Russell, Michael (2017) *Top quark physics in the Large Hadron Collider era.* PhD thesis.

Top Quark Physics in
the Large Hadron Collider era

Michael Russell

Particle Physics Theory Group,
School of Physics & Astronomy,
University of Glasgow

July 2017

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy

Abstract

We explore various aspects of top quark phenomenology at the Large Hadron Collider and proposed future machines. After summarising the role of the top quark in the Standard Model (and some of its well-known extensions), we discuss the formulation of the Standard Model as a low energy effective theory. We isolate the sector of this effective theory that pertains to the top quark and that can be probed with top observables at hadron colliders, and present a global fit of this sector to currently available data from the LHC and Tevatron. Various directions for future improvement are sketched, including analysing the potential of boosted observables and future colliders, and we highlight the importance of using complementary information from different colliders. Interpretational issues related to the validity of the effective field theory formulation are elucidated throughout. Finally, we present an application of artificial neural network algorithms to identifying highly-boosted top quark events at the LHC, and comment on further refinements of our analysis that can be made.

Acknowledgements

First and foremost I must thank my supervisors, Chris White and Christoph Englert, for their endless support, inspiration and encouragement throughout my PhD. They always gave me enough freedom to mature as a researcher, whilst providing the occasional necessary nudge to keep me on the right track. I also have to thank David Miller for his friendly advice and irreverent sense of humour, and for helping me settle into the group in Glasgow, and Christine Davies, for financial support for several research trips, and for creating a wonderful place to do physics.

This thesis would not have been written without the foundations laid down by Liam Moore, so I have to thank him for the countless hours he has put into the project, and for being a mate. Of the other students in Glasgow, I also have to mention Karl Nordström for many useful conversations, both about the work presented here and often tangential (but always illuminating) topics, and his occasional computational wizardry. In Heidelberg, special mention must go to Torben Schell, for his patience in showing me some of the many ropes of jet substructure, and to Tilman Plehn for stimulating collaboration, and for helping me start the next chapter in my life as a physicist. I thank the students and postdocs in both of these places for making such a friendly working environment.

Throughout my PhD I have been fortunate to work with several excellent experimentalists; James Ferrando and Andy Buckley in Glasgow, and Gregor Kasieczka in Zürich. I thank them for all they have taught me about collider physics (and especially Andy for coding assistance in the rocky early days of TopFitter), and I hope to collaborate with them again in the future. Though I never had the opportunity to work with Sarah Boutle or Chris Pollard directly, their unwavering fondness for a pint helped keep me sane after many a long day.

Finally, I have to acknowledge the huge intellectual (and general) debt I owe to my parents. Though they do not share my passion for physics, they have always supported me along the way, and politely endured far too many wearisome physics rants to mention. This thesis is dedicated to them. Thanks also to Daniel and Lucy for trying to keep me in the real world.

Declaration

I declare that this thesis is the result of my own original work and has not been previously presented for a degree. In cases where the work of others is presented, appropriate citations are used. Chapters 1 and 2 serve as an introduction to the research topics presented in the rest of the thesis. Chapters 3 to 5 are a result of my own original work, in collaboration with the authors listed below.

Work appearing in this thesis was presented in the following publications [1–5]:

- A. Buckley, C. Englert, J. Ferrando, D. J. Miller, L. Moore, M. Russell, and C. D. White, "Global fit of top quark effective theory to data," *Phys. Rev.* **D92** no. 9, (2015) 091501, arXiv:1506.08845 [hep-ph].

- A. Buckley, C. Englert, J. Ferrando, D. J. Miller, L. Moore, M. Russell, and C. D. White, "Constraining top quark effective theory in the LHC Run II era," *JHEP* **04** (2016) 015, arXiv:1512.03360 [hep-ph].

- C. Englert, L. Moore, K. Nordström, and M. Russell, "Giving top quark effective operators a boost," *Phys. Lett.* **B763** (2016) 9–15, arXiv:1607.04304 [hep-ph].

- A. Buckley, C. Englert, J. Ferrando, D. J. Miller, L. Moore, K. Nordström, M. Russell, and C. D. White, "Results from TopFitter," in *9th International Workshop on the CKM Unitarity Triangle (CKM 2016) Mumbai, India, November 28-December 3, 2016.* 2016. arXiv:1612.02294 [hep-ph]. http://inspirehep.net/record/1502116/files/arXiv:1612.02294.pdf.

- G. Kasieczka, T. Plehn, M. Russell, and T. Schell, "Deep-learning Top Taggers or The End of QCD?," *JHEP* **05** (2017) 006, arXiv:1701.08784 [hep-ph].

- C. Englert and M. Russell, "Top quark electroweak couplings at future lepton colliders," arXiv:1704.01782 [hep-ph].

My specific contributions to these chapters are as follows. In chapter 3, I set up the EFT fitting framework, and used the MATHEMATICA model developed by Liam Moore to simulate events, construct theory observables and derive confidence limits on the operators considered. In chapter 4, I wrote the boosted analysis and derived the various projected limits on the operators considered. All numerical results in those chapters were derived by myself. In chapter 5, my contributions involved generating events, writing the analysis code to construct the fat jets and subsequent images used, formulating the boosted decision tree used as a comparison, and helping to design the neural network architecture used. All figures in this thesis were generated by the author, with the following exceptions: Fig. 3.2 (A. Buckley), Figs. 4.13-4.14 (C. Englert), Fig. 5.5 and Figs. 5.7 -5.9 (T. Schell).

Contents

A historical introduction

The origins of what is now known as the 'Standard Model' of particle physics can be traced to the late 1940s. The attention of theoretical physicists was centred on how to consistently embed the postulates of quantum mechanics; the laws describing (sub)atomic particles, within the framework of special relativity; the laws of motion for objects with very large velocities. Their efforts culminated in the development of quantum electrodynamics; the quantum theory of the electromagnetic field. The problems associated with the infinities arising in calculations had been brought under control by the development of renormalisation, which allowed properties of the electron and photon to be calculated to high precision, showing extraordinary agreement with experiment.

It had already been known for some time, however, that electrodynamics could not be the full story. Firstly, it was immediately obvious that the stability of the atomic nucleus would not withstand the electrostatic repulsion between the positively-charged protons, therefore an additional force must have been present to stabilise the nucleus. This force had to be strong, and extremely short-ranged (no further than the typical size of a nucleus) so was dubbed the *strong interaction*. Moreover, the observation of certain types of radioactive decay, which necessitated the existence of a new, extremely light particle (what is now called the neutrino), could not be accommodated with the known facts about electromagnetism and the strong force. These interactions did not allow for processes which changed electric charge, which nuclear β-decay plainly did. Due to the relatively long lifetimes associated with these decay processes, the force responsible was called the *weak interaction*.

The first attempt to write down a theory of the strong nuclear force was made by Yukawa in 1935 [6]. He proposed that the force binding together the nucleus was due to an interaction between protons and neutrons mediated by a scalar particle he dubbed the 'pion', which he calculated should have a mass of around one tenth of the proton mass. Tentative discoveries of these pions were made in 1947 in photographic emulsion recordings of cosmic ray showers [7]. The problem with the theory was that it could not predict anything precisely. The strong force is (by definition) strong, and all the known calculational tools of the day relied on treating the interactions as small perturbations, and the particles as almost non-interacting. These approximations failed spectacularly when applied to the strong force.

Besides, the cosmic ray observations posed an additional problem: in Yukawa's model the protons, neutrons and mediating pions were considered fundamental; that is, not containing any substructure. However, experiments made using the more advanced bubble chamber discovered a slew of new particles - similar in properties to the pion and proton, but with different masses. By Yukawa's token, each of these new particles were just as

fundamental as the proton or the pion. By the mid 1950s, however, dozens of such particles had been discovered, none of which had, or could have, been predicted, and with no underlying theory to relate them. Fundamental physics in this era was in a state of excited disarray.

Motivated purely by the observed properties of nuclear β-decay, and by the discovery of the neutron by Chadwick two years previously [8], Fermi wrote down the first model of the weak interaction in 1934 [9], which modelled β-decay as a contact interaction in which a neutron decays into a proton by emitting an electron and a neutrino. It was extremely successful in predicting observables in β-decay such as the electron energy spectrum, but soon it was realised that the theory was non-renormalisable: the infinities that had plagued early calculations in QED cropped up again, but unlike in QED, they could not be removed. Therefore it was abandoned as a fundamental theory.

An improvement came in the form of the intermediate vector boson model [10–15], where, rather than a contact interaction, the decay was described as mediated by the exchange of vector bosons, completely analogously to the pions that mediated the strong force in Yukawa's theory, and the photons of QED. This immediately raised the problem that, unlike the photon and pion, these vector bosons had not been discovered, and had to be extremely heavy to give the correct radioactive decay rates. More startlingly, new tests of the properties of β-decay showed that the weak interaction, unlike all the other known forces, violated parity symmetry [16,17], i.e. it was able to distinguish between left and right. Any new theory of the weak interaction would have to be radically different in structure to accommodate these facts.

Despite these puzzles, progress was made in the 1950s and early 1960s on two fronts. The first was the quark model of Gell-Mann and Zweig [18,19]. In an effort to classify the myriad of new particles emerging from the bubble chamber experiments, they postulated that rather than being fundamental, these particles were composed of smaller particles (Gell-Mann coined them 'quarks', in a literary homage to James Joyce's *Finnegan's Wake*). Requiring only 3 *flavours* of quark (which were dubbed the *up*, *down* and *strange*) as elements of the global symmetry group SU(3), that is, location-independent transformations on the quark fields by 3×3 unitary matrices, the model was able to accommodate the observed mass spectrum of many of the observed mesons and baryons, and predicted new ones, several of which were duly discovered. Still, for most physicists, these quarks were no more than an idealisation, a useful bookkeeping device for classifying the bubble chamber results, and few took their existence seriously as fundamental particles.

The other major development was made in 1954 by Yang and Mills [20] who made the observation that the electromagnetic interaction could be described as resulting from a U(1) *gauge* symmetry, a type of local symmetry where the fields receive a location-dependent phase transformation but the full theory is left invariant. They observed that

since the proton and neutron were almost equal in mass, it was instructive to model them as elements of a 2D symmetry group, hence they modelled the proton-neutron nuclear force as originating from a local SU(2) symmetry, which they dubbed *isospin*. The immediate difference from QED was that the gauge boson of this force would be self-interacting, unlike the photon. This was perfectly allowed by the experimental facets of the strong interaction, and offered intriguing insights into the possibility of constructing a quantum theory of gravity.

The bugbear of the gauge theories suggested by Yang and Mills was how to accommodate mass. If gauge symmetry was to be an exact symmetry of the weak and strong interactions, the mediating particles of this symmetry; the gauge bosons, would have to be massless and so (it was presumed), these forces would have to be long-ranged, just like in electromagnetism, which they clearly were not. On the other hand, adding in mass terms for the gauge bosons violated the gauge symmetry explicitly, defying the point of introducing it in the first place, and, as was later shown [21, 22], just as in the case of the intermediate vector boson models, led to unacceptable physical behaviour when extrapolated to higher energies. Despite their mathematical beauty, the application of gauge theories to particle physics was stymied in the late 1950s and early 1960s by the apparent incompatibility between the symmetry patterns of the theory and the basic observation of particle masses.

The missing piece of the puzzle emerged from a completely different area of physics: superconductivity. When a conductor is cooled below a certain ultra-cool temperature, it displays the bizarre property of having almost no electrical resistance. Anderson noted [23] that this effect could be explained by the photons which transmit the electric forces inside the bulk of the superconductor effectively gaining a mass, which would break the long range electromagnetic interactions and allow currents to flow with effectively zero resistance. The U(1) electromagnetic symmetry still remained in the conductor, but it was *spontaneously*, rather than explicitly, broken when it entered the superconducting phase. Anderson speculated that this phenomena might have important consequences for the application of gauge theories to elementary particle physics.

Ideas about spontaneously broken symmetries were already being tested in particle physics, but in the wrong way. It had been suggested by several authors that the known approximate symmetries of the strong interactions, such as the proton-nucleon isospin and the symmetries of Gell-Mann's quark model, could have originated from the spontaneous breaking of some exact symmetry of the system, perhaps at a higher energy. This idea suffered an apparently fatal blow when it was proved by Goldstone, Salam and Weinberg [24, 25] that any spontaneously broken symmetry would necessarily lead to the appearance of massless, interacting scalar bosons. These scalars would have easily been observed experimentally long ago and they had not, and, ignorant of the developments in

superconductivity, most particle theorists regarded Goldstone's theorem as the death knell for this idea.

The importance of Anderson's observations was immediately appreciated by Higgs, however. He had been trying to find a loophole in Goldstone's theorem, and showed that if the symmetry of the system was not global but local [26](as in the gauge theories of Yang and Mills), the unwanted massless scalars that resulted from the symmetry breaking would be absorbed by the gauge bosons, giving them a mass. In this way, gauge symmetry could be preserved whilst still giving masses to the gauge bosons, just as the U(1) electromagnetic gauge symmetry was preserved while the photons inside the superconductor gained an effective mass. The massive Yang-Mills problem had, in principle, been solved. He speculated that this mechanism could be applied to a gauge theory of the weak interaction, and would allow the vector bosons mediating the interaction to gain the mass they needed. Crucially, he also predicted the appearance of a new, massive scalar boson, which now bears his name [27]. The same ideas were published at almost exactly the same time by Brout and Englert [28], and by Guralnik, Hagen and Kibble [29], who were attempting to give a mass to the pion in a gauge theory of the strong interaction.

The ideas of Brout, Englert, Guralnik, Hagen, Higgs and Kibble were put to use by Weinberg [30] and Salam [31] in 1967. They wrote down a gauge group with an SU(2)×U(1) symmetry, as had been suggested in studies by Glashow [32] and Salam and Ward in 1961 [33]. Unlike the earlier papers, however, which contained explicit gauge boson mass terms, they applied the Higgs mechanism to it, and showed that the charged vector bosons mediating the weak interaction (the W bosons) gained a very large mass, at least forty times the mass of the proton. Echoing Higgs' conclusions, they predicted the appearance of a massive scalar. They also predicted the appearance of a heavier still, electrically neutral vector boson. Since this was the last new particle required by the model, it was called the Z. After the symmetry breaking, an unbroken subgroup remained, this was identified as electromagnetism, with a massless photon. Hence, their theory unified the weak and electromagnetic interactions into one single model. This 'electroweak' theory, along with the Higgs mechanism, forms one of the two pillars of what we now call the Standard Model of elementary particles.

Still, the spectre of renormalisability loomed over the electroweak theory. Though renormalisability had been demonstrated in electromagnetism by Feynman, Schwinger, Tomonaga and Dyson twenty years earlier [34–44], little progress had been made in tackling the problem for the more sophisticated Yang-Mills theories. The question then remained of whether the symmetry breaking mechanism spoiled the renormalisability of the electroweak theory, in which case it would have been little improvement over its Fermi and intermediate vector boson model predecessors. Most theorists thought the answer to this question was yes, and so the unified electroweak theory received little attention at first. In a series

of *tour de force* calculations [45], it was shown by 't Hooft and Veltman that the gauge theories of Yang and Mills were, in fact, renormalisable. First this was demonstrated in the massless case [46], and then in the more involved case where the vector bosons get their masses from the Higgs mechanism [47]. After this work was published, interest in the Weinberg-Salam-Glashow model exploded, and a dedicated experimental program for uncovering the precise gauge structure of the electroweak interaction took off.

The other half of the Standard Model is the strong interaction. In contrast to the agitation that engulfed the theory community in the 1950s and early 1960s, particle accelerators evolved rapidly during this era. These developments allowed the quark model of Gell-Mann to be put to a crucial test in experiments at SLAC and MIT in 1967 [48, 49]. The experiments drew analogy with the famous Rutherford experiment of 1912, in which a beam of α-particles were fired at a strip of gold foil, and the rare collisions in which the α-particles rebounded from the foil provided evidence for a positively charged nucleus within the centre of the atom. The SLAC-MIT experiments instead fired a high-energy beam of electrons into a fixed proton target, to probe the putative inner structure of the proton. The results were unequivocal. High scattering rates at large angles were observed, containing events with the detection of the scattered electron and large numbers of hadrons. The sole explanation for these events was that the electron shattered the proton into its intermediate pieces, which interacted some time later to re-form into into the various types of hadrons observed in the final state.

These results gave strong weight to the existence of quarks as fundamental particles, but that relied on a strange presumption: while the scattering rates at large angles were consistent with electromagnetic interactions between the electrons and the proton's sub-components, they could only be explained if the strong interaction between the proton's inner parts was much weaker at high-energy, i.e. that when the incoming electron approached the proton, it saw the proton constituents as almost non-interacting. Then when the constituents became separated again after the collision, the strong interaction between them switched on again, binding them into the hadrons that were observed in the final state. These ideas were put on a firm mathematical footing in the parton model of Feynman and Bjorken [50, 51]. Still, the behaviour of the strong force: weak at short distances and strong at long distances (like the restoring force on a stretched elastic band) was at odds with all the known forces at the time. Electromagnetism and gravity both become weaker as the separation between the interacting objects is increased.

A few years later, in 1973, Gross, Politzer and Wilczek [52, 53] discovered a class of theory which exhibited precisely this property, which is known as *asymptotic freedom*. They were exactly the same theories used to construct the electroweak interaction: Yang-Mills gauge theories. The strong interaction was modelled by an SU(3) gauge symmetry between the quarks. Their models assumed that the quarks had, as well as electric charge, a

property of 'strong' charge, which came to be known as colour charge. The SU(3) structure presumed there were 3 types of such colour. The theory thus became known as quantum chromodynamics, in analogy with QED 30 years previously. The gauge bosons of this interaction were ultimately responsible for binding together the quarks into hadrons such as the proton, so they were dubbed gluons.

The puzzle remained of how the strong interaction remained short-ranged, and how come the massless gluons had not been observed themselves. It was initially presumed that the SU(3) symmetry was broken so that the gluons gained mass, as in the case of the Ws and Zs, but were too heavy to observe. It was soon realised that the gluons could indeed be massless, but the same phenomena which bound the quarks together was responsible for keeping the gluons confined with the hadrons. This phenomena of *confinement* has intriguing consequences for how we view mass; most of the mass of hadrons such as the proton and neutron (and, by extension, most of the mass of the observable Universe) originates not from the mass of their constituents, but in the binding energy between the constituents. An analytic proof of confinement in Yang-Mills theory has not been rigorously obtained, and the Clay Mathematical Foundation continues to offer a \$1 million prize for a first-principles solution. Nonetheless, numerical studies on the lattice have demonstrated that confinement is indeed a property of QCD.

The Standard Model was beginning to take shape, but several observations were still at odds with its early predictions. Firstly, the observed rates of certain types of *strangeness violating* weak decay processes were much lower than expected from the electroweak model. Secondly, it had been known since the 1950s that the weak interactions violated the so-called CP-symmetry, effectively a symmetry between matter and antimatter, by a small amount, but the electroweak model contained no terms which violated CP. The first problem was solved by Glashow, Ilopoulos and Maiani [54], who showed that the existence of a fourth flavour of quark (they dubbed it the *charm*) was able to suppress the rates by much more than the naïve prediction. It was quickly realised that this fourth quark could lead to many new different kinds of meson. The simplest of these would be a bound state of a charm quark and antiquark. This was promptly discovered in 1974 [55, 56], with a production rate and mass in excellent agreement with the Standard Model predictions. It was then realised by Kobayashi and Maskawa [57], building on earlier work by Cabbibo [58],, that CP-violation could be obtained by adding in a 3rd generation (a fifth and sixth flavour) of quarks, these were called the *top* and the *bottom*. The bottom quark was discovered in 1978 [59,60], the much heavier top quark in 1995 [61,62]. For theoretical reasons mainly pertaining to the cancellation of anomalies, it was also presumed that a 3rd generation of leptons would exist, a heavier extension of the electron and muon. The charged lepton of this generation: the τ, was discovered in experiments between 1974 and 1977 [63], its neutrino was finally discovered in 2000 [64].

	Fermions			Bosons
Quarks	$\begin{pmatrix} u \\ d \end{pmatrix}$	$\begin{pmatrix} c \\ s \end{pmatrix}$	$\begin{pmatrix} t \\ b \end{pmatrix}$	γ g
Leptons	$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix}$	$\begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix}$	$\begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}$	$W^\pm$ Z^0 H

Table 0.1: *Particle content of the Standard Model of particle physics. The mass of each matter generation increases from left to right. The fermions form doublets which differ in electric charge by 1: for the quarks this is of the form $(q_u, q_d) = (+2/3e, -1/3e)$, for the leptons this is $(q_l, q_\nu) = (-1e, 0)$.*

These discoveries complete what we now call the Standard Model of particle physics: The matter content consists of six quarks and six leptons, in three generations of increasing mass. The force carries are the gauge bosons: the photon of electromagnetism, the W and Z of the weak interaction, and the gluon mediating the strong interaction. Underpinning all of it is the Higgs boson, which breaks the electroweak symmetry and gives mass to the W and Z, and also to the quarks and charged leptons. This is summarised in table 0.1. The model is strikingly minimal; with just a handful of particles it can explain all the observable matter content in the Universe, and its interactions (other than gravity). The main concern is the *ad hoc* nature of its structure. It was largely cobbled together to fit experiment, and all the parameters relating to the masses of the fermions, the mixing between different generations, and the relative strengths of the weak, electromagnetic and strong interactions, as well as the mass of the Higgs boson, are not predicted by it, and have to be determined by experiment.

Whatever aesthetic qualms one may have about its structure, however, the successes of the Standard Model as a physical theory describing Nature have been nothing short of astounding. The first coup of the Glashow-Weinberg-Salam model was the discovery of the neutral currents in 1973 [65, 66], lending strong indirect evidence for the existence of the Z. The W and Z bosons were discovered outright at CERN in 1983 [67–70], four years after Glashow, Weinberg and Salam were awarded the Nobel Prize in Physics for their electroweak theory. The gluon was discovered in three-jet events at the PETRA experiment in 1978 [71]. The predictions of the Standard Model continued to be tested throughout the 1980s in collider experiments across the world. These efforts culminated in the precision electroweak measurements at LEP and SLC [72], which probed the SU(3)×SU(2)×U(1) gauge structure to per-mille level accuracy, providing incontrovertible evidence that the Standard Model is an excellent description of Nature up to energies around 100 GeV. The last outstanding piece of the theory; the Higgs boson, was discovered in 2012 [73, 74], 48 years after it was first hypothesised, and a detailed program for the precise measurements of its properties is now well underway [75].

Despite these triumphs, ever since the inception of the Standard Model, physicists have been looking for evidence for new physics which will take us beyond the current SM paradigm. After the unification of the weak and electromagnetic interactions into one simple model, it was natural to ask if the strong interaction could be unified with the electroweak in a single gauge group. More ambitious still were attempts to include gravitational interactions in such a framework, a so-called theory of everything. Early attempts at these models pointed out that the unification would happen at a very large energy scale, inaccessible to any conceivable future collider experiment. However, several pieces of indirect evidence point to new physics just above the electroweak scale, well within reach of current colliders.

With a centre-of-mass energy of 14 TeV, the LHC is best poised to answer the question of whether new physics beyond the SM resides at the TeV energy scale. However, there are a large number of well-motivated scenarios, and their experimental signatures are often very similar. Given the huge catalogue of measurements published by the LHC, and the possibility of different manifestations of new physics hiding in many of them, it is best to ask not "Does my new physics model explain this particular measurement better than the Standard Model alone?" but "Which consistent theory best describes *all* the data?". This has led to renewed interest in being able to describe the data in a model-independent way. Effective field theory provides such a description.

Since its discovery in 1995, the top quark continues to mystify physicists with its properties. As it is the only fermion with a mass around the electoweak scale ($m_t = 173$ GeV), and as the precise mechanism which breaks the electroweak symmetry is unexplained in the Standard Model, the top quark usually plays a special role in theories of physics beyond the Standard Model. The top quark sector is thus one of the many well-motivated places to look for the effects of potential new physics, as only now are its properties beginning to be scrutinised with high precision. The language of effective field theory provides a powerful, systematic way of doing this. This is the primary topic of this thesis.

The thesis is structured as follows. In chapter 1, I will discuss the unique role of the top quark in the Standard Model. In chapter 2, I will outline some of the main hints for physics beyond the Standard Model and some well-studied new physics models, and their relevance for top quark phenomenology. I will then describe the formulation of the Standard Model as a low energy effective theory where all the ultraviolet degrees of freedom have been integrated out, and the sector of this effective theory that can be studied with top measurements from hadron colliders. In chapter 3 I will discuss a global fit of the top quark sector of the Standard Model EFT to data from the LHC. Chapter 4 is concerned with refinements of the analysis of chapter 3, such as how the increase in LHC energy from 8 to 13 TeV can be best exploited; how 'boosted observables' that draw on high-momentum transfer final states can improve the fit results, and how proposed future lepton colliders

can complement results extracted from hadron collider measurements. In chapter 5 we move away from effective theory, and study how the performance of certain algorithms for reconstructing 'boosted' final states may be augmented by recent developments in machine learning, before summarising the conclusions of this thesis.

The motivations for this work are thus threefold: 1. With the abundant data from the LHC, top properties can be examined with precision for the first time. 2. The top quark continues to be a sensible place to search for new physics. 3. Effective field theories are a powerful tool for confronting new physics models with data in a systematic way. It is worth remembering that the story of the Standard Model began with an effective theory when Fermi wrote down his model of nuclear β-decay in 1935. Perhaps it would be fitting if the story ended with one as well.

1 The top quark in the Standard Model

1.1 Introduction

The top quark was discovered in 1995 by the CDF [61] and DØ [62] experiments at the Tevatron. Still, only recently have its couplings begun to be measured with sub-10% level accuracy, thanks to the much higher production rates at the LHC and the large integrated luminosity collected over the total lifetime of the Tevatron. The role that the top quark might play in specific realisations of electroweak symmetry breaking is just beginning to be tested. Before we can turn to these questions, however, we must summarise the unique role of the top quark within the Standard Model. This is the subject of this chapter.

This introductory chapter is structured as follows. In section 1.2 I discuss the building blocks of the Standard Model of particle physics: the unified $SU(2)_L \times U(1)_Y$ electroweak theory; the $SU(3)_C$ gauge theory of the strong interaction known as quantum chromodynamics and the role of spontaneously broken local gauge symmetry via the Higgs mechanism, before discussing the free parameters of the Standard Model. In section 1.3 I discuss some generalities about hadron collider phenomenology, including the main theoretical uncertainties that crop up in scattering calculations. In section 1.4 I discuss the main production mechanisms for top quarks at hadron colliders, and some properties of top production and decay, before summarising in section 1.5.

1.2 The Standard Model of Particle Physics

The Standard Model has three main ingredients. Firstly, there is quantum chromodynamics (QCD) [52, 53, 76]: the theory of the strong interaction between 'coloured' quarks and gluons (the mediators of this interaction), described by a gauge group with a local $SU(3)_C$ symmetry. Secondly, the electroweak theory described by the model of Glashow, Salam and Weinberg [30–32], which unifies the electromagnetic and weak interactions of quarks and leptons under the gauge group $SU(2)_L \times U(1)_Y$; its charges are weak isospin L and weak hypercharge Y. Finally, there is the celebrated Higgs mechanism [27–29]: a complex scalar field doublet (with four degrees of freedom) whose potential acquires a non-zero minimum which spontaneously breaks the electroweak symmetry into a $U(1)$ group describing QED; its charge is the familiar electromagnetic coupling. Three of the four degrees of freedom form the longitudinal polarisation states of the $W^\pm$ and Z^0 bosons, giving mass to these particles and thus being responsible for the phenomena of nuclear β-decay and other weak processes. The remaining one forms a massive scalar particle: the Higgs boson. The Higgs mechanism is also responsible for giving mass to the quarks and charged leptons through a Yukawa-type interaction.

1.2.1 Before electroweak symmetry breaking

The Standard Model before electroweak symmetry breaking has two types of field:

Matter fields ψ**:** Since the weak interactions are known to violate parity, the matter fields are constructed out of left-handed and right-handed (*chiral*) fermions.

$$\psi = \psi_L + \psi_R \tag{1.1}$$

where

$$
\begin{aligned}
\psi_L &= P_L \psi \quad \text{with} \quad P_L = \frac{(1 - \gamma_5)}{2} \\
\psi_R &= P_R \psi \quad \text{with} \quad P_R = \frac{(1 + \gamma_5)}{2}.
\end{aligned}
\tag{1.2}
$$

The operators $P_{L,R}$ project out the chiral states of each fermion, whose kinetic terms in the Dirac Lagrangian $\mathcal{L}_{\text{Dirac}}$ can thus be decomposed:

$$\bar{\psi} \gamma^\mu \partial_\mu \psi = \bar{\psi}_L \gamma^\mu \partial_\mu \psi_L + \bar{\psi}_R \gamma^\mu \partial_\mu \psi_R \tag{1.3}$$

i.e. massless fermions decouple into chiral components. For the SM we have three generations of left-handed and right-handed spin-$\frac{1}{2}$ fermions which can be categorised into quarks and leptons. To reproduce the chiral structure of the weak interaction, left-handed fermions are in weak-isospin doublets, and right-handed fermions fall into weak-isospin singlets.

$$
\begin{aligned}
Q_1 &= \begin{pmatrix} u \\ d \end{pmatrix}_L, & u_{R_1} &= u_R, & d_{R_1} &= d_R, & L_1 &= \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, & e_{R_1} &= e_R^- \\
Q_2 &= \begin{pmatrix} c \\ s \end{pmatrix}_L, & u_{R_2} &= c_R, & d_{R_2} &= s_R, & L_2 &= \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, & e_{R_2} &= \mu_R^- \\
Q_3 &= \begin{pmatrix} t \\ b \end{pmatrix}_L, & u_{R_3} &= t_R, & d_{R_3} &= b_R, & L_3 &= \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L, & e_{R_3} &= \tau_R^-.
\end{aligned}
\tag{1.4}
$$

Members of each doublet have 3rd component of weak isospin $I_f^3 = \pm\frac{1}{2}$, which is related to U(1)$_Y$ hypercharge Y_f and electric charge Q_f by

$$Y_f = Q_f - I_f^3, \tag{1.5}$$

where

$$Y_{L_i} = -\frac{1}{2}, \quad Y_{e_R} = -1, \quad Y_{Q_i} = +\frac{1}{6}, \quad Y_{u_{R_i}} = +\frac{2}{3}, \quad Y_{d_{R_i}} = -\frac{1}{3}. \tag{1.6}$$

These hypercharge assignments ensure the fermions have the correct electric charge: isodoublets differ in electric charge by 1: for the quarks this is of the form $(q_u, q_d) = (+2/3e, -1/3e)$, for the leptons this is $(q_l, q_\nu) = (-1e, 0)$. The quark fields are charged under $\mathrm{SU}(3)_C$, i.e. each quark appears as a triplet of 3 colours, whereas the leptons are singlets. This important feature ensures that the anomaly cancellation condition

$$\sum_f Y_f = 0, \tag{1.7}$$

where the sum runs over all fermions in a generation, is satisfied. Hence gauge-invariance is not spoiled by radiative corrections and the theory remains renormalisable.

Gauge fields V_μ: These are the spin-1 bosons that mediate the electroweak and strong interactions. The $\mathrm{SU}(2)_L$ symmetry of the electroweak sector gives rise to 3 vector fields $W_\mu^{1,2,3}$ corresponding to the generators T^I ($I = 1, 2, 3$), expressed in terms of the Pauli matrices τ^I as*

$$T^I = \frac{1}{2}\tau^I \quad ; \quad \tau^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \tau^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \tau^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{1.8}$$

which satisfy the commutation relations

$$[T^I, T^J] = i\epsilon^{IJK}T^K, \tag{1.9}$$

where ϵ^{IJK} is the antisymmetric tensor. The $\mathrm{U}(1)_Y$ symmetry corresponds to a vector field B_μ, which has the unique generator Y. The strong sector has an $\mathrm{SU}(3)_C$ symmetry, corresponding to 8 gluon fields $G_\mu^{1,...,8}$, expressed in terms of the 3×3 Gell-Mann matrices T^A which we do not write explicitly here and which satisfy

$$[T^A, T^B] = if^{ABC}T^C \quad \text{and} \quad \mathrm{Tr}[T^A T^B] = \frac{1}{2}\delta^{AB}, \tag{1.10}$$

where f^{ABC} denote the $\mathrm{SU}(3)_C$ structure constants. From these fields one may construct gauge invariant field strength tensors

$$\begin{aligned} G_{\mu\nu}^A &= \partial_\mu G_\nu^A - \partial_\nu G_\mu^A + g_s f^{ABC} G_\mu^B G_\nu^C \\ W_{\mu\nu}^I &= \partial_\mu W_\nu^I - \partial_\nu W_\mu^I + g\epsilon^{IJK} W_\mu^J W_\nu^K \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu, \end{aligned} \tag{1.11}$$

where g_s and g respectively denote the $\mathrm{SU}(3)_C$ and $\mathrm{SU}(2)_L$ coupling constants. The $\mathrm{U}(1)_Y$

---

*Throughout this thesis, weak isospin indices are denoted by lowercase Roman letters $\{i, j, k \ldots\} \in \{1, 2\}$, while $\mathrm{SU}(3)_C$ and $\mathrm{SU}(2)_L$ adjoint indices are denoted by uppercase Roman: $\{A, B, C \ldots\} \in \{1 \ldots 8\}$ and $\{I, J, K \ldots\} \in \{1, 2, 3\}$.

coupling is denoted as g'.

To couple the matter fields to the gauge fields, we replace the ordinary derivative ∂_μ with the gauge covariant derivative D_μ:

$$D_\mu = \partial_\mu + ig_s T^A G_\mu^A + igT^I W_\mu^I + ig'\frac{Y}{2}B_\mu. \tag{1.12}$$

which leads to matter-gauge couplings of the form

$$\text{fermion-gauge couplings} \quad : \quad g_i \bar{\psi} V_\mu \gamma^\mu \psi \tag{1.13}$$

In addition to matter-gauge interactions, the non-Abelian nature of the SM leads to self-interactions among the gauge bosons, which we can generically class into 3-point and 4-point couplings:

$$\begin{aligned} \text{3-point couplings} \quad &: \quad -g_i \text{Tr}(\partial_\mu V_\nu - \partial_\nu V_\mu)[V_\mu, V_\nu] \\ \text{4-point couplings} \quad &: \quad g_i^2 \text{Tr}[V_\mu, V_\nu]^2 \end{aligned} \tag{1.14}$$

where $g_i \in \{g_s, g, g'\}$. The Standard Model Lagrangian at this point consists only of kinetic terms for massless fermions and gauge bosons:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermion}}, \tag{1.15}$$

where

$$\begin{aligned} \mathcal{L}_{\text{gauge}} &= -\frac{1}{4}G_{\mu\nu}^A G^{A,\mu\nu} - \frac{1}{4}W_{\mu\nu}^I W^{I,\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \\ \mathcal{L}_{\text{fermion}} &= i\bar{L}_i D_\mu \gamma^\mu L_i + i\bar{e}_{Ri} D_\mu \gamma^\mu e_{Ri} + i\bar{Q}_i D_\mu \gamma^\mu Q_i + i\bar{u}_{Ri} D_\mu \gamma^\mu u_{Ri} + i\bar{d}_{Ri} D_\mu \gamma^\mu d_{Ri}. \end{aligned} \tag{1.16}$$

It is manifestly invariant (by construction) under local $\text{SU}(3)_C \times \text{SU}(2)_L \times \text{U}(1)_Y$ gauge transformations. For instance, under an $\text{SU}(2)_L$ transformation,

$$\begin{aligned} e_{Ri} &\to e'_{Ri} = e_{Ri} \\ L_i &\to L'_i = e^{-i\omega^I(x)T^I} L_i. \end{aligned} \tag{1.17}$$

So we see that the $\text{SU}(2)_L$ singlets ψ_R are trivially $\text{SU}(2)_L$ invariant and therefore do not couple to the corresponding gauge fields $W_\mu^{1,2,3}$. So far, the theory is self-consistent. When we try to include particle masses, however, we run into two problems:

1. **Fermion Masses:** Explicit fermion masses take the form $\mathcal{L}_{\text{mass}} = -m\bar{\psi}\psi$, which
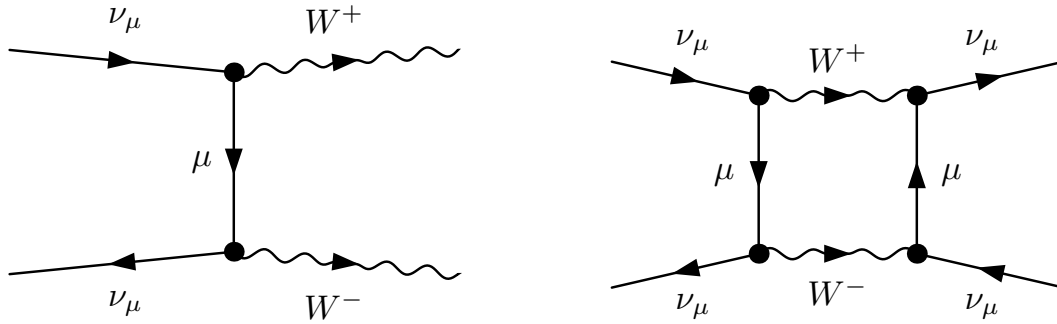
Figure 1.1: *Left: The tree-level process $\nu_\mu \nu_\mu \to W^+ W^-$ in the intermediate vector boson model, which violates unitarity in the high-energy limit. Right: One-loop neutrino scattering in the IVB model; a non-renormalisable interaction.*

when decomposed into chiral components become:

$$m\bar{\psi}\psi = m\bar{\psi}\left(\frac{(1-\gamma_5)}{2} + \frac{(1+\gamma_5)}{2}\right)\psi = m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R) \qquad (1.18)$$

which is not $SU(2)_L$ invariant, as it mixes left-handed and right-handed fermion components.

2. **Gauge boson masses:** The observed short-range of the weak interaction ~ 0.1 fm tells us that the vector bosons mediating to the weak interaction have masses of order ~ 10 GeV. However, when we include explicit mass terms in the Lagrangian, it is easy to see they are not gauge invariant. Using the simple U(1) case of QED with a massive photon as an example:

$$\frac{1}{2}M_A^2 A_\mu A^\mu \to \frac{1}{2}M_A^2 (A_\mu - \frac{1}{e}\partial_\mu \alpha)(A_\mu - \frac{1}{e}\partial^\mu \alpha) \neq \frac{1}{2}M_A^2 A_\mu A^\mu \qquad (1.19)$$

To appreciate the problems caused by explicit breaking of gauge invariance, consider the propagator for a generic massive vector boson.

$$\frac{i}{p^2 - M^2}\left(-g^{\mu\nu} + \frac{p^\mu p^\nu}{M^2}\right) \qquad (1.20)$$

and the weak-interaction process $\nu_\mu \bar{\nu}_\mu \to W^+ W^-$, the leading order Feynman diagram for which is sketched on the left-hand side of Fig. 1.1.

Although this process would be rather difficult to implement experimentally, its cross-section can be straightforwardly calculated. In the high-energy limit, the $p^\mu p^\nu / M^2$ term, corresponding to the longitudinal W polarisation states, will dominate contributions to the

cross-section. In fact, one finds for this process

$$\sigma \sim G_F^2 E^2 \qquad (1.21)$$

where G_F is a coupling constant which must have dimensions of $(\text{energy})^{-2}$: i.e. the cross-section grows quadratically with energy. We can decompose the scattering amplitude A for this process into partial waves f_ℓ of orbital angular momentum ℓ:

$$A = 16\pi \sum_{\ell=0}^{\infty} (2\ell + 1) P_\ell(\cos\theta) f_\ell \qquad (1.22)$$

where P_ℓ are the Legendre polynomials and θ is the scattering angle. Noting that for $2 \to 2$ processes with massless external legs the cross-section is given by $d\sigma/d\Omega = |A|^2/64\pi^2 s$, with $d\Omega = 2\pi d\cos\theta$, the total cross-section is then

$$
\begin{aligned}
\sigma &= \frac{8\pi}{s} \sum_{\ell=0}^{\infty} \sum_{\ell'=0}^{\infty} (2\ell + 1)(2\ell' + 1) f_\ell f_{\ell'} \int_{-1}^{1} d\cos\theta\, P_\ell(\cos\theta) P_{\ell'}(\cos\theta) \\
&= \frac{16\pi}{s} \sum_{\ell=0}^{\infty} (2\ell + 1)|f_\ell|^2
\end{aligned}
\qquad (1.23)
$$

where the orthogonality condition $\int d\cos\theta\, P_\ell P_{\ell'} = 2\delta_{\ell\ell'}/2l + 1$ was used. From the optical theorem (a simple consequence of unitarity), σ is equal to the imaginary part of the forward ($\theta = 0$) scattering amplitude [77], so that, at each order in the partial wave expansion, unitarity requires:

$$
\begin{aligned}
|f_\ell|^2 = \text{Im}(f_\ell) &\Rightarrow [\text{Re}(f_\ell)]^2 + [\text{Im}(f_\ell)]^2 = \text{Im}(f_\ell) \\
&\Rightarrow [\text{Re}(f_\ell)]^2 + [\text{Im}(f_\ell) - \frac{1}{2}]^2 = \frac{1}{4}
\end{aligned}
\qquad (1.24)
$$

which is just the equation of a circle in the $[\text{Re}(f_\ell), \text{Im}(f_\ell)]$ plane, of radius $\frac{1}{2}$ centred at $[0, \frac{1}{2}]$. Hence $|\text{Im}(f_\ell)| \leq 1$, and the cross-section in each partial wave projection has the unitarity bound

$$\sigma \leq \frac{16\pi(2\ell + 1)}{s}. \qquad (1.25)$$

Comparing Eq. (1.25) with Eq. (1.21), we see that unitarity is violated at some finite energy. Plugging the numbers in we find this is around $E \sim 1$ TeV [78–80], indicating that beyond this energy the theory is perturbatively not well-defined.

Since this is only a perturbative statement, one might well argue that the theory may still be consistent if strong dynamics take over in this regime. However, we could instead consider the case where the W bosons appear as virtual particles, e.g. in the one-loop
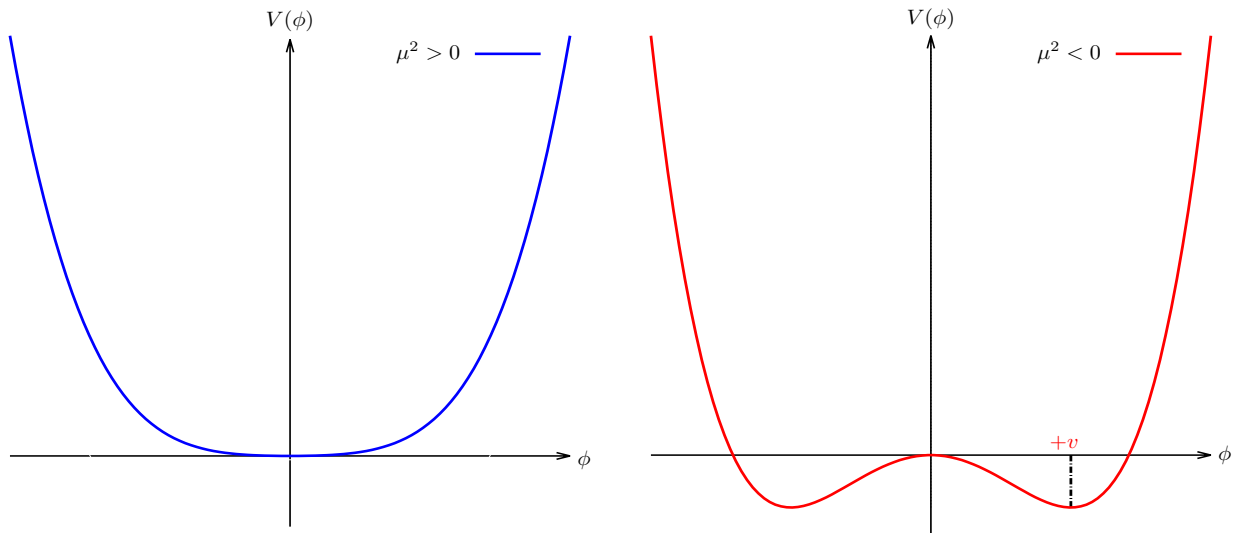
Figure 1.2: *The potential $V(\phi)$ of Eq. (1.26) in the cases $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right).*

process $\nu_\mu \bar{\nu}_\mu \rightarrow \nu_\mu \bar{\nu}_\mu$, as depicted on the right-hand side of Fig. 1.1, in which the longitudinal states W_L lead to quadratically divergent loop-momenta. Renormalising this divergence would require the inclusion of a counterterm corresponding to a four-neutrino vertex. However, no such vertex exists in the theory. Hence, the theory is non-renormalisable, and cannot be expected to make predictions for arbitrarily high-energies.

To summarise, it seems there is a fundamental conflict between constructing renormalisable gauge theories for particle physics, and allowing particles in those theories to have mass. If there was a way to generate mass *dynamically*, i.e. not through explicit mass terms but through a gauge-invariant interaction between fields, perhaps the gauge principle can be saved. The Higgs mechanism provides such an interaction.

1.2.2 The Higgs mechanism

As a warmup, we consider the example of a real scalar field ϕ with the Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi) \quad \text{where} \quad V(\phi) = \frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4 \tag{1.26}$$

$\mathcal{L}$ is invariant under reflections $\phi \rightarrow -\phi$. For $\mathcal{L}$ to describe any physical system, λ must be positive-semidefinite, otherwise the potential is unbounded from below. μ^2 can take positive or negative values, however. For $\mu^2 > 0$ the minimum of the potential (in quantum field theoretic terms, its *vacuum expectation value* $\langle 0|\phi|0\rangle$) is located at the origin $\phi_0 = 0$. In this case $\mathcal{L}$ is just the Lagrangian of a spin-zero particle of mass μ, as shown in the left-hand side of Fig. 1.2. However, when $\mu^2 < 0$ this no longer represents the Lagrangian

of a particle of mass μ. The minima of the potential are now at

$$\langle 0|\phi|0\rangle = \phi_0 = \pm\sqrt{-\frac{\mu^2}{\lambda}} \equiv \pm v \qquad (1.27)$$

The field has picked up a non-zero vacuum expectation value v, as highlighted in the right-hand side of Fig 1.2. To extract the interactions of this theory, we expand the field around $\phi = v$. Defining $\phi = v + \sigma$, $\mathcal{L}$ is, up to constant terms

$$\mathcal{L} = \frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma + \mu^2\sigma^2 - \sqrt{-\mu^2\lambda}\sigma^3 - \frac{1}{4}\lambda^4 \qquad (1.28)$$

The theory now describes a new scalar field of mass $m_\sigma^2 = -2\mu^2$, with trilinear and quartic self-interactions. The σ^3 term breaks the original reflection symmetry; that is, a symmetry of the Lagrangian is no longer a symmetry of the vacuum, it has been *spontaneously broken*.

The next simplest example of spontaneously broken symmetry is that of four scalar fields (equivalently a complex doublet of scalars) with Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi_i\partial^\mu\phi_i - \frac{1}{2}\mu^2\phi_i\phi_i + \frac{1}{4}\lambda(\phi_i\phi_i)^2 \qquad (1.29)$$

which is invariant under the transformation $\phi_i = R_{ij}\phi_j$ where R_{ij} are 4-dimensional orthogonal matrices, i.e. transformations under the rotation group in four dimensions, O(4). Setting $\mu^2 < 0$ and expanding around the minima at $\phi_i = (0,0,0,v)$, where $v^2 = \frac{\mu^2}{\lambda}$, $\mathcal{L}$ becomes

$$\begin{aligned}
\mathcal{L} = &\frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma + \mu^2\sigma^2 - \sqrt{-\mu^2\lambda}\sigma^3 - \frac{1}{4}\lambda^4 \\
&+ \frac{1}{2}\partial_\mu\pi_i\partial^\mu\pi_i - \frac{1}{4}\lambda(\pi_i\pi_i)^2 - \lambda v\pi_i\pi_i\sigma - \frac{1}{2}\pi_i\pi_i\sigma^2,
\end{aligned} \qquad (1.30)$$

where i now runs from 1 to 3, and $\sigma = \phi_4 - v$, $\pi_i = \phi_i$. Again, a massive σ boson with mass $m_\sigma^2 = -2\mu^2$ has appeared, but so have three massless pions, among which there is a residual O(3) symmetry. This is an example of a general property of spontaneously broken continuous symmetries known as Goldstone's theorem [25], which can be stated as follows:

For a continuous symmetry group $\mathbb{G}$ spontaneously broken down to a subgroup $\mathbb{H}$, the number of broken generators is equal to the number of massless scalars that appear in the theory.

The O(N) group has $N(N-1)/2$ generators, so O($N-1$) has $(N-1)(N-2)/2$ and $N-1$ Goldstone bosons appear (the above example is the case of $N=4$).

The above example applied to global symmetries, but if the mechanism is extendable

to local (gauge) symmetries, it would provide a viable way of giving mass to the vector bosons of the weak interaction. We begin with the case of an Abelian U(1) symmetry, with the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\phi)^*(D^\mu\phi) - V(\phi) \quad \text{where} \quad V(\phi) = \mu^2\phi^*\phi + \lambda(\phi^*\phi)^2, \qquad (1.31)$$

where $D_\mu = \partial_\mu + igA_\mu$ is the usual covariant derivative. This is invariant under local U(1) transformations:

$$\phi(x) \to e^{i\theta(x)}\phi(x), \quad A_\mu \to A_\mu + \frac{1}{g}\partial_\mu\theta(x). \qquad (1.32)$$

The case $\mu^2 > 0$ corresponds to scalar QED: interactions between a charged scalar of mass μ and a massless vector boson, with an additional four-point scalar self-interaction. For $\mu^2 < 0$, ϕ as usual obtains a non-zero vev, and the potential is minimised at

$$\langle 0|\phi|0\rangle = \sqrt{-\frac{\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{2}}. \qquad (1.33)$$

Expanding the potential around the vev,

$$\phi(x) = \frac{1}{\sqrt{2}}(v + \phi_1(x) + i\phi_2(x)) \equiv \frac{1}{\sqrt{2}}(v + H(x))e^{i\chi(x)/v}, \qquad (1.34)$$

the Lagrangian describing the vacuum state is now

$$\begin{aligned}
\mathcal{L} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (\partial_\mu - igA_\mu)\phi^*(\partial^\mu + igA^\mu)\phi - \mu^2\phi^*\phi - \lambda(\phi^*\phi)^2 \\
&= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}(\partial_\mu\phi_1)^2 + \frac{1}{2}(\partial_\mu\phi_2)^2 + \mu^2\phi_1^2 + \frac{1}{2}g^2v^2A_\mu A^\mu + gvA_\mu\partial^\mu\chi \\
&\quad + (\text{interaction terms}).
\end{aligned} \qquad (1.35)$$

The photon has obtained a mass $M_A^2 = g^2v^2$, the scalar particle ϕ_1 has a mass $M_{\phi_1}^2 = -2\mu^2$. The ϕ_2 has apparently emerged as the Goldstone boson of this symmetry breaking.

However, $\mathcal{L}$ now contains the bilinear term $gvA_\mu\partial^\mu\chi$, which neither corresponds to an interaction or a field strength. The symmetry breaking has also apparently created an extra degree of freedom. Before, there were four: two in the massless photon and two in the complex field ϕ. Now there appear to be five: three for the massive photon, and one each for ϕ_1 and ϕ_2. The resolution of this paradox lies in the fact that we are free to make a gauge transformation:

$$\phi(x) \to e^{-i\chi(x)/v}\phi(x), \quad A_\mu \to A_\mu + \frac{1}{gv}\partial_\mu\chi(x) \qquad (1.36)$$

which removes all $\chi(x)$ terms from the Lagrangian. Counting degrees of freedom, we see the

massless photon has absorbed the Goldstone boson, and gained mass: it has a longitudinal polarisation state. The U(1) symmetry has been spontaneously broken, leading to a massive vector boson and the appearance of a massive scalar boson. This is the Higgs mechanism.

1.2.3 The Higgs mechanism in the Standard Model

To apply the Higgs mechanism to the Standard Model, we need to generate mass for the $W^{\pm}$ and Z^0 bosons, whilst keeping the photon massless. So the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ electroweak symmetry should be broken to a U(1) subgroup describing electromagnetism. This means that at least 3 degrees of freedom are needed. We also want to introduce a gauge-invariant interaction that gives masses to fermions without mixing chiral components. The simplest object that satisfies these criteria is an SU(2) doublet of scalar fields ϕ

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{1.37}$$

where the superscript denotes the electric charge in each component. We add the usual ϕ^4 Lagrangian $\mathcal{L}_{\mathrm{Higgs}}$ to the SM Lagrangian in Eq. (1.16)

$$\mathcal{L}_{\mathrm{Higgs}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \quad \text{where} \quad V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \tag{1.38}$$

$V(\Phi)$ gets a minimum at $\Phi^\dagger \Phi = \mu^2/2\lambda$, which we take to be in the neutral direction to preserve $\mathrm{U}(1)_{e.m}$

$$\langle \Phi \rangle_0 = \langle 0|\Phi|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{1.39}$$

with $v = \sqrt{-\frac{\mu^2}{\lambda}}$. Expanding around the vev as before:

$$\Phi = \frac{e^{i\theta^I(x)\tau^I/v}}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{1.40}$$

and expanding out the covariant derivative term in $\mathcal{L}_{\mathrm{Higgs}}$, we have

$$
\begin{aligned}
(D_\mu \Phi)^\dagger (D^\mu \Phi) &= \left| \left(\partial_\mu + ig\frac{\tau^I}{2}W_\mu^I + ig'\frac{Y}{2}B_\mu \right) \right|^2 \\
&= \frac{1}{2} \left| \begin{pmatrix} \partial_\mu + \frac{i}{2}(gW_\mu^3 + g'\frac{Y}{2}B_\mu) & i\frac{g}{2}(W_\mu^1 - iW_\mu^2) \\ i\frac{g}{2}(W_\mu^1 + iW_\mu^2) & \partial_\mu - \frac{i}{2}(gW_\mu^3 - g'\frac{Y}{2}B_\mu) \end{pmatrix} \begin{pmatrix} 0 \\ v + h \end{pmatrix} \right|^2 \\
&= \frac{1}{2}(\partial_\mu h)^2 + \frac{1}{8}(v+h)^2|W_\mu^1 + iW_\mu^2|^2 + \frac{1}{8}(v+h)^2|gW_\mu^3 - g'B_\mu|^2 \\
&\quad + (\text{interaction terms}).
\end{aligned} \tag{1.41}
$$

Eq. (1.41) shows that there are terms mixing the fields W_μ^3 and B_μ. The physical bosons must be superpositions of these fields such that there are no mixing terms. The physical fields can be obtained by performing the rotation

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}, \tag{1.42}$$

where the weak mixing/Weinberg angle

$$\tan\theta_W \equiv \frac{g'}{g}, \tag{1.43}$$

has been introduced. With this, Eq. (1.41) becomes

$$(D_\mu\Phi)^\dagger(D^\mu\Phi) = \frac{1}{2}(\partial_\mu h)^2 + \frac{g^2v^2}{4}W_\mu^+W^{-\mu} + \frac{g^2v^2}{8\cos^2\theta_W}Z_\mu Z^\mu + 0A_\mu A^\mu, \tag{1.44}$$

where $W^\pm = (W^1 \mp W^2)/\sqrt{2}$. The W and Z bosons have acquired masses

$$M_W = \frac{1}{2}gv, \quad M_Z = \frac{1}{2}\frac{gv}{\cos\theta_W} \tag{1.45}$$

i.e. there is a mass relation

$$M_Z = \rho M_W \cos\theta. \tag{1.46}$$

The parameter ρ has been introduced: at tree-level $\rho = 1$ but radiative quantum effects give corrections to this relation. The SU(2) gauge structure of the electroweak theory ensures that these corrections are small, however; a feature known as *custodial symmetry* [81]. Different choices of representations for the Higgs field (e.g. an SU(2) triplet) would not protect the ρ parameter from large corrections. The linear combination A has remained massless, so is to be identified with the photon. To see that a U(1) subgroup remains unbroken, consider the symmetry associated with the generator

$$Q \equiv T^3 + \frac{Y}{2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{1.47}$$

where we have included the explicit representation of Y as a 2×2 identity matrix. Then

$$Q|0\rangle \sim \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ v+h \end{pmatrix} = 0 \tag{1.48}$$

i.e. the symmetry associated with this generator is unbroken by the vacuum, so the corresponding field $gW^3 + g'B \equiv A$ is massless. We can similarly expand the potential terms

29

around the vacuum:

$$V(\Phi) = \frac{\mu^2}{2}(0, v + h)\begin{pmatrix} 0 \\ v + h \end{pmatrix} + \frac{\lambda}{4}\left| (0, v + h)\begin{pmatrix} 0 \\ v + h \end{pmatrix}\right|^2$$

$$= -\lambda^2 v^2 - \lambda v h^3 - \frac{\lambda}{4}h^4 + (\text{constants}).$$

(1.49)

So the scalar particle has gained a mass $m_h^2 = -2\mu^2 = 2\lambda v^2$, and has trilinear and quartic self-interactions. This is the Higgs boson. Next we turn to the issue of generating masses for the fermions. This too can be done in a gauge invariant way through a Yukawa-type interaction $\phi\bar{\psi}\psi$

$$\mathcal{L}_{\text{Yukawa}} = -y_e\bar{L}\Phi e_R - y_d\bar{Q}\Phi d_R - y_u\bar{Q}\tilde{\Phi}u_R + h.c.$$

(1.50)

where $\tilde{\Phi} = i\tau_2\Phi^*$ is used instead of Φ for the up quark because the vev is in the lower component of the Higgs doublet. Upon spontaneous symmetry breaking, we have, e.g. for the electron

$$\mathcal{L}_{\text{Yukawa}} = -\frac{1}{\sqrt{2}}y_e\;(\bar{\nu}_L \quad \bar{e}_L)\begin{pmatrix} 0 \\ v + h \end{pmatrix}e_R + h.c. \Rightarrow -\frac{y_e v}{\sqrt{2}}\bar{e}e + \text{interaction term}, \quad (1.51)$$

and similarly for the up and down quarks. To summarise, using an SU(2) doublet Φ we have generated masses for both the W and Z vector bosons and the fermions. The SU(2)$_L \times$ U(1)$_Y$ symmetry is no longer apparent in the vacuum; it has been spontaneously broken down to an unbroken U(1) subgroup, identified as electromagnetism. The color SU(3) symmetry is also unbroken, so has been omitted in this section. Because gauge invariance has not been explicitly broken, the Standard Model remains renormalisable [45, 46] and unitary [82, 83] up to high energies. The Standard Model can thus be summarised by the following Lagrangian[†].

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{Yukawa}} + \mathcal{L}_{\text{Higgs}}$$

(1.52)

where

$$\begin{aligned}
\mathcal{L}_{\text{gauge}} &= -\frac{1}{4}G_{\mu\nu}^A G^{A,\mu\nu} - \frac{1}{4}W_{\mu\nu}^I W^{I,\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \\
\mathcal{L}_{\text{fermion}} &= i\bar{L}_i D_\mu\gamma^\mu L_i + i\bar{e}_{Ri}D_\mu\gamma^\mu e_{Ri} + i\bar{Q}_i D_\mu\gamma^\mu Q_i \\
&\quad + i\bar{u}_{Ri}D_\mu\gamma^\mu u_{Ri} + i\bar{d}_{Ri}D_\mu\gamma^\mu d_{Ri} \\
\mathcal{L}_{\text{Yukawa}} &= -y_e\bar{L}\Phi e_R - y_d\bar{Q}\Phi d_R - y_u\bar{Q}\tilde{\Phi}u_R + h.c. \\
\mathcal{L}_{\text{Higgs}} &= (D_\mu\Phi)^\dagger(D^\mu\Phi) - \mu^2\Phi^\dagger\Phi - \lambda(\Phi^\dagger\Phi)^2.
\end{aligned}$$

(1.53)

[†]The full Lagrangian also contains gauge-fixing and Fadeev-Popov 'ghost' terms to eliminate redundant gauge field configurations. For brevity these are not included here.

1.2.4 The parameters of the Standard Model

For one generation of fermions, the free parameters in the Standard Model are:

- The three gauge couplings $\{g_s, g, g'\}$

- The two parameters in the Higgs potential $V(\phi)$: μ and λ

- The three Yukawa coupling constants $\{y_u, y_d, y_e\}$

Although these are the 'fundamental' parameters, they are typically expressed in terms of the more directly measurable quantities:

$$
\begin{aligned}
\tan\theta_W &= \frac{g'}{g} \\
e &= g\sin\theta_W \\
m_H &= \sqrt{2}\mu = \sqrt{2\lambda}v \\
M_W &= \frac{g\mu}{2\sqrt{\lambda}} = \frac{gv}{2} \\
m_f &= \frac{y_f\mu}{\sqrt{2\lambda}} = \frac{y_f}{\sqrt{2}}v.
\end{aligned}
\tag{1.54}
$$

Once these parameters have been measured precisely, predictions for M_Z and G_F (the Fermi coupling) can be made. Thus the interaction strengths of the entire electroweak sector of the Standard Model are fixed by seven parameters (the strong interaction is determined by 1: g_s) .

Adding additional generations brings some complications, however. For instance, the presence of a second and third generation of quarks leads to the Yukawa couplings

$$
- [y_d]_{ij}\bar{Q}_i\Phi d_{Rj} - [y_u]_{ij}\bar{Q}_i\Phi u_{Rj} + h.c.
\tag{1.55}
$$

where i, j are generation/flavour indices. The Yukawa couplings are now 3×3 matrices, and off-diagonal terms are perfectly allowed by gauge-invariance. This would mix quarks of different flavour. To obtain the physical particles we diagonalise the mass matrix and extract the terms bilinear in each field, just as we did to extract the physical Z and A fields. This can be done by performing a unitary rotation on each quark field. However, this means that we must also rotate the quark kinetic terms, so the off-diagonal structure has merely been transferred to the fermion-gauge couplings. To relate the weak eigenstates to the mass eigenstates, the convention is to define the up-type quarks as in the mass-eigenstate basis to begin with, then to relate the down-type quark weak eigenstates q' to the mass eigenstates q through a unitary rotation

$$(d' \quad s' \quad b') \; = \mathbf{V} \; (d \quad s \quad b) \tag{1.56}$$

where $\mathbf{V}$ is the 3×3 Cabibbo-Kobayashi-Maskawa (CKM) matrix [57, 58], whose values are [84]:

$$
\begin{aligned}
\mathbf{V_{CKM}} &= \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \\
&= \begin{pmatrix} 0.97427 \pm 0.00014 & 0.22536 \pm 0.00061 & 0.00355 \pm 0.00015 \\ 0.22522 \pm 0.00061 & 0.97343 \pm 0.00015 & 0.0414 \pm 0.0012 \\ 0.00886 \pm 0.00033 & 0.0405 \pm 0.0012 & 0.99914 \pm 0.00005 \end{pmatrix}
\end{aligned} \tag{1.57}
$$

To count the parameters of this matrix, we first note that a general unitary 3×3 matrix has nine independent parameters. With six quarks we can absorb five relative phases into the quark field strengths $q \rightarrow e^{i\theta}q$, which leaves four independent parameters: three mixing angles (akin to the Euler rotation angles) and a residual complex phase. The off-diagonal terms in the CKM matrix are subleading, and a well-known parametrisation of the CKM matrix which mimics this structure is due to Wolfenstein [85], which can be approximated as:

$$\mathbf{V_{CKM}} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & 1 - A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{1.58}$$

where the complete expression involving $\{\bar{\rho}, \bar{\eta}\} = \{\rho, \eta\}(1 - \lambda^2/2 + \mathcal{O}(\lambda^4))$ has not been displayed here. For massless neutrinos, there is no analogous mixing in the lepton sector: the weak eigenstates are the same as the mass eigenstates, a property of the Standard Model known as *lepton universality*.

Hence for the Standard Model with three generations, we have the following free parameters, 18 in total[‡].

- The 8 parameters mentioned above.

- Three extra Yukawa couplings for each additional generation: six in total.

- Four parameters in the CKM matrix: $\{A, \bar{\rho}, \lambda, \bar{\eta}\}$.

[‡]Adding in neutrino masses would bring in 7 more parameters: 3 masses and 4 PMNS [86–88] mixing angles. For the purposes of this thesis neutrinos can be considered massless.

32

The most up-to-date values of these parameters[§], expressed through more conveniently measurable quantities, are shown in Table 1.1 [84].

Parameter	Value	Parameter	Value
α	1/137.035 999 074(44)	m_e	0.510 998 928(11) MeV
α_s	0.1185(6)	m_μ	105.6583715(35) MeV
G_F	$1.166\ 378\ 7(6) \times 10^{-5}$ GeV^{-2}	m_τ	1776.82(16) MeV
m_h	125.7(4) GeV	m_u	$2.3^{+0.6}_{-0.5}$ MeV
M_W	80.385(15) GeV	m_d	$4.8^{+0.5}_{-0.3}$ MeV
A	$0.814^{+0.023}_{-0.024}$	m_c	1.275(25) GeV
λ	0.22537(61)	m_s	95(5) MeV
$\bar{\rho}$	0.117(21)	m_t	$173.21 \pm 0.51 \pm 0.71$ GeV
$\bar{\eta}$	0.353(13)	m_b	4.18(3) GeV

Table 1.1: *The 18 free parameters of the Standard Model with massless neutrinos.*

A few remarks are in order:

- The coupling constants α are related to their respective gauge coupling parameters by

$$\alpha = \frac{e^2}{4\pi} \quad , \quad \alpha_s = \frac{g_s^2}{4\pi} \tag{1.59}$$

Coupling constant is something of a misnomer, however. The renormalisation of the couplings by higher-order corrections ensures the values of these parameters depend on the scale at which they are resolved. This will be discussed in more detail later. In Table 1.1 α is quoted at the scale $Q^2 = 0$, whereas α_s is quoted at $Q^2 = M_Z^2$.

- In a similar way, the quark masses $m_i = \frac{y_i v}{\sqrt{2}}$ are renormalised by QCD effects (QED renormalisation of lepton masses is negligible), and the values quoted refer to the 'running masses' in the $\overline{MS}$ renormalisation scheme, each evaluated at the scale $\mu = 2$ GeV, with the exception of the top quark.

- The top quark presents an additional ambiguity: the measured value quoted above is obtained from fitting Monte Carlo templates with different input values for the top quark mass. This was formerly interpreted as equal to the top quark *pole* mass: the renormalised mass corresponding to the pole in the top propagator. This analogy is flawed, however, due to subtleties in the showering and hadronisation of partons in the Monte Carlo. The ambiguity of definition here introduces a theoretical uncertainty of ~ 1 GeV in additional to the statistical and systematic errors quoted above. This issue is discussed in more detail in the next section.

[§]There is also a 19th parameter: the QCD θ-term, which does not lead to physical effects in perturbation theory, but can be generated by non-perturbative instanton effects. This will be discussed in some detail in chapter 2.

The predictions of the Standard Model have been tested in numerous fixed-target and collider experiments over the last forty years. The rich phenomenology of the strong interaction has been extensively studied in electron-positron collisions and deep-inelastic scattering in electron-proton events at HERA [89]. In addition, the precision measurements carried out at LEP, the Tevatron and elsewhere [72, 90] have probed the electroweak couplings to sub-percent level accuracy. The 2012 discovery [73, 74] of a Higgs boson at a mass of ~ 125 GeV by the ATLAS and CMS experiments has filled in the last piece of the SM picture, and studies of the Higgs sector to a similar precision are now underway [75].

The subject of this thesis concerns the properties of the top quark, and how they may be probed at hadron colliders, so the next section reviews some general features of hadron collider machines such as the LHC and Tevatron, before reviewing the physics of the top quark that may be studied with them.

1.3 Hadron collider physics

1.3.1 Scattering theory

The starting point for calculating scattering amplitudes in quantum field theory is the S-matrix, which can be split into a trivial 'free-propagation piece' and a scattering piece T.

$$S = \mathbb{1} + iT = \mathbb{1} + i\delta^4(p_f - p_i)\mathcal{M}_{fi} \tag{1.60}$$

The delta function appears in every scattering amplitude to enforce momentum conservation, so can be factored out of any scattering amplitude to define the matrix-element $\mathcal{M}_{fi}$. From this, and Fermi's golden rule, the cross-section for producing a general final state X from initial state particles a_1 and a_2 of momenta p_1 and p_2 is

$$\sigma(a_1(p_1)a_2(p_2) \rightarrow X) = \frac{1}{\Phi} \int d\Pi_n |\mathcal{M}_{fi}|^2 \tag{1.61}$$

where

$$\Phi = |v_{a_1} - v_{a_2}|(2E_{a_1})(2E_{a_2}) \tag{1.62}$$

is the flux factor, defined in terms of the relative velocities of the incoming beams in the lab frame, v_{a_i}, and

$$\int d\Pi_n = \int \prod_{i=1}^{n} \frac{d^3k_i}{(2\pi)^3 2E_{k_i}} = \int d\text{LIPS}(X) \tag{1.63}$$

is the volume of the n-body final-state Lorentz invariant phase-space. Cross-sections can be calculated order-by-order in perturbation theory, provided that coupling constants are not too large so that higher-order terms in the perturbation series can be neglected. The initial states relevant for hadron collisions are the constituents of the hadron: quarks and
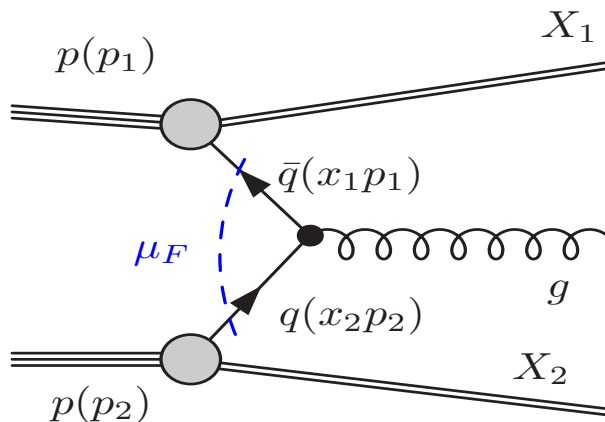
Figure 1.3: *Schematic diagram for the process $pp \rightarrow gX$, where X denotes any other final state products of the collision. The factorisation scale μ_F separates the perturbative 'hard' process from the non-perturbative parton densities of the incoming hadrons.*

gluons. The perturbative cross-section is calculated completely in terms of quarks and gluons (collectively known as 'partons') and is related to the full hadronic cross-section by

$$\sigma(p(k_1)p(k_2) \rightarrow X) = \sum_{i,j} \int_0^1 dx_1 dx_2 f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \hat{\sigma}_{ij}(x_1, x_2, s, \alpha_S(\mu_R, Q^2)). \quad (1.64)$$

Here $\hat{\sigma}_{ij \rightarrow X}$ is the partonic cross-section for final state X from partons i and j, where $i, j \in \{q, g\}$. f_i are the parton density functions (pdfs): the probability of finding a parton i in the proton with fraction x of the total proton momentum. To obtain the full hadronic cross-section, we calculate the partonic cross-section for momenta $x_1 p_1$ and $x_2 p_2$, then integrate these over the full range of x for each proton, then sum over all allowed partonic subprocesses. This is illustrated in Fig. 1.3.

The parameters $\mu_{R,F}$ are arbitrary scales necessarily introduced in fixed-order perturbation theory. Calculations for hadron colliders are plagued by theoretical uncertainties, which can be broadly classed into three categories: scale dependence, pdf uncertainties and finite accuracies for Standard Model parameters such as α_s, which enter as inputs into the calculation. Here we briefly discuss each of these in turn.

1.3.2 Scale uncertainties

It is well-known that quantities calculated beyond the leading Born approximation in quantum field theory often feature ultraviolet divergences. These arise from quantum fluctuations with unconstrained high-momenta. For a certain class of quantum field theories, it is possible to remove these divergences by defining the theory at some renormalisation scale μ_R which separates the low energy field theory from the unknown short-distance
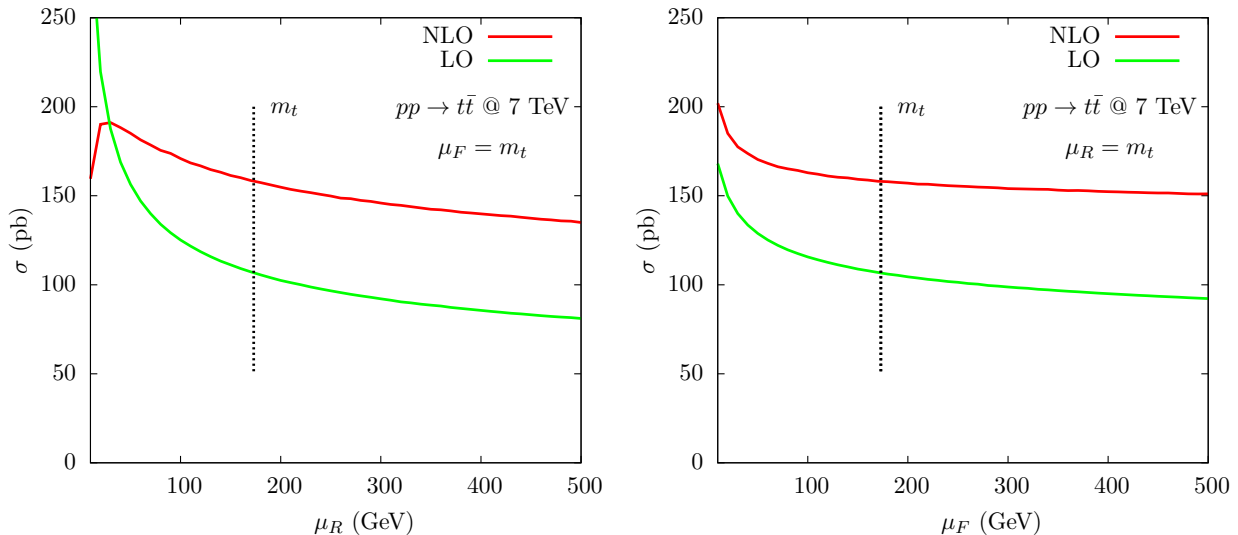
Figure 1.4: *Renormalisation (left) and factorisation (right) scale dependence of the $pp \to t\bar{t}$ cross-section at 7 TeV, using the MSTW2008 pdf sets.*

physics and allows one to make low-energy predictions regardless of the underlying degrees of freedom. Although separated out, the degrees of freedom at different scales have the effect of introducing a *scale dependence* of the coupling constants and masses of the theory. In QED, for instance, the effective electromagnetic coupling runs from $\alpha = 1/137$ at the scale $\mu_R = 0$ to $\alpha(\mu_R = M_Z) \sim 1/129$.

The renormalisation scale is an arbitrary parameter and so predictions for physical quantities should be independent of μ_R. The renormalisation group equations define precisely how the renormalised couplings should vary with scale such that order-by-order in perturbation theory, measurable quantities are independent of μ_R. Truncating the perturbative expansion at a fixed order, however, means that the cancellation of μ_R in physical quantities is incomplete, i.e. there is a residual dependence on μ_R proportional to the next order in the perturbative expansion.

It is not immediately clear which value of μ_R should be chosen for a process, but it should be a characteristic energy scale entering the process that absorbs the large logarithms $\log(s/\mu_R^2)$ arising from separate scales involved in the process. To estimate the size of unknown higher-order corrections, one typically varies the scale over the range $\mu/2 \leq \mu_R \leq 2\mu$, using the variation in the prediction as the scale uncertainty. For most processes where NNLO corrections have been calculated, they have been found to lie in the scale uncertainty band of the NLO estimate, which vindicates this rather *ad hoc* procedure. Counter-examples exist, however, where the NLO and NNLO scale uncertainty bands do not overlap, such as in cases where widely different scales enter[¶], and to truly quantify higher-order effects there is no substitute for doing the actual calculation.

[¶]One example is associated Higgs production $pp \to HV$ [91].

One must also choose the factorisation scale for a process. This defines at what energy we separate the hard (high-momentum) scattering process cross-section, which is calculated in perturbation theory, from the parton density functions (PDFs) for the incoming protons, which are extracted from data.

The factorisation scale is also not a physical quantity, it is a definition of what energy scale corresponds to the partonic process and what falls into the definition of the incoming protons. Any initial state radiation with energy $E < \mu_F$ is absorbed into the hadron. Again, there is no 'correct' scale, one simply chooses a value typical of the process and varies over $[1/2,2]$ to estimate the uncertainty. For most of our predictions (save a few special cases) we set a common central scale $\mu_R = \mu_F = \mu = m_t$ and vary both independently over $[\mu/2,2\mu]$. The dependence of the total $t\bar{t}$ production cross-section at the 7 TeV LHC, at leading and next-to-leading order, on these scales is sketched in Fig. 1.4, showing that the range $[\mu/2,2\mu]$ captures most of the scale dependence.

1.3.3 The parton densities

The inner structure of the proton is determined by quantum chromodynamics in the strongly coupled, low-momentum transfer regime where perturbative techniques are not valid, so the parton densities f_i are not calculable from first principles. The choice of pdf set introduces an additional theoretical uncertainty and several such sets are available. Of course, predictions should be independent of the set used: the structure of the proton at a certain energy scale is a universal physical property. In practice, however, the different approaches each group uses to extract parton densities from data introduce systematic uncertainties leading to different results.

Due to the vastly different methodologies used by the main pdf groups, and the different input measurements used in their fits, it is often not possible to compare their results in an unbiased way. Instead the discrepancies resulting from different pdf choices are resolved in the most conservative way, by calculating predictions for each of the main pdf groups: CT14 [92], MMHT [93] and NNPDF [94], and taking the maximum range as an additional ('pdf') uncertainty. This prescription is the recommendation of the PDF4LHC [95] working group, and is the one adopted throughout this thesis, unless otherwise stated.

1.3.4 Standard Model parameters

In addition to the theoretical uncertainties arising from scale and pdf choices, an additional source arises from the finite precision with which the SM parameters entering the calculation have been measured. Of the 18 parameters in Table 1.1, the two most relevant for top quark production are the strong coupling constant α_s and the top quark mass m_t.
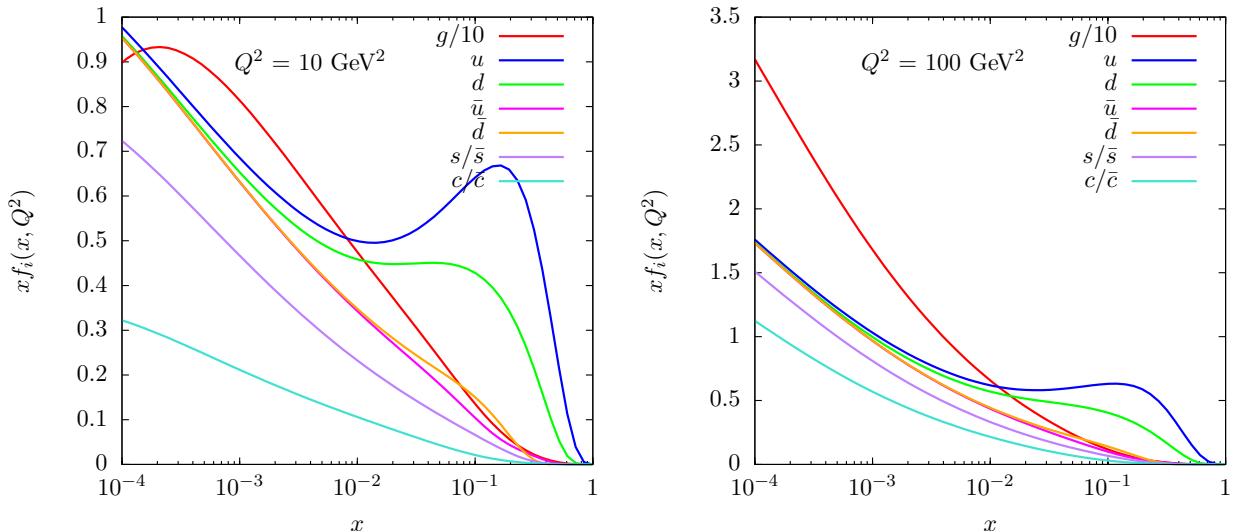
Figure 1.5: *Parton distribution functions for the proton from the MSTW2008 NNLO fit [96], at low (left) and high (right) momentum transfer. Uncertainties are not shown.*

The former is known to sub per-mille accuracy, having been extracted mainly from high-precision e^+e^- experiments at LEP and SLC, as well as through deep inelastic scattering measurements. Its value, quoted at the Z-pole, can be calculated at any energy scale using the QCD β-function, which has recently been calculated to five-loop accuracy [97–99]. It is therefore one of the most precisely measured quantities of the Standard Model, and the inclusion or omission of its experimental uncertainty rarely has a substantial effect on perturbative QCD predictions.
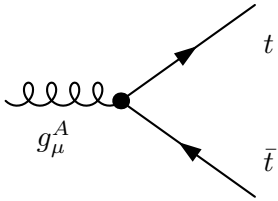
The top quark mass, however, presents additional challenges, as mentioned above. Experimentally, the top quark mass has been measured to sub percent-level accuracy, This is typically achieved by directly reconstructing the top quark from its decay products: a b-tagged jet and either a charged lepton and missing transverse energy, or an additional pair of jets. Kinematic distributions of these decay products are constructed, e.g. the reconstructed top mass m_t, and Monte-Carlo predictions ('templates') with different values of m_t^{MC} are fit to the data. The best-fit value is then defined as the top quark mass. This is a well-defined statistical procedure. However, ambiguity arises when relating m_t^{MC} to a renormalised mass in quantum field theory.

The top quark mass is renormalised by self-energy corrections. The UV divergent pieces of these corrections are absorbed unambiguously into the running of the mass. However, different treatments of the finite corrections admit different definitions of what is meant by a 'mass' in quantum field theory. The most intuitive is the *pole* mass, the mass corresponding to the pole in the propagator, where *all* divergent and finite corrections are absorbed into the mass. Owing to non-perturbativity, however, loop corrections with momenta $\lesssim 1$ GeV (the QCD hadronisation scale) cannot be calculated, which defines a maximum precision on the pole mass definition. A Monte Carlo generator never runs into such problems. The MC

top mass is defined as the pole in the hard matrix element. When this is interfaced to the parton shower, which generates successive parton splittings at increasingly low momenta, self-energy corrections are ignored, so they must be viewed as already included in the definition of m_t^{MC}. However, when the typical parton momenta in the shower reaches $\mathcal{O}(1$ GeV), showering stops and the hadronisation model takes over. There is thus a fundamental precision of ~ 1 GeV with which we can relate m_t^{pole} to the experimentally measured m_t^{MC}, and this uncertainty should be included in any calculations involving m_t (see Ref. [100] for a recent review of these issues).
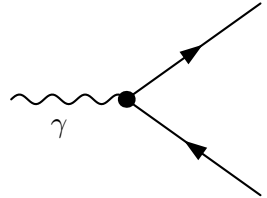
1.4 Top quark physics at hadron colliders

The top quark couples directly to all of the Standard Model gauge and Higgs bosons. The interaction with gluons is described by a vectorial fermion-gauge coupling $\bar{\psi}\psi A_\mu$

$$= -ig_s T^A \gamma^\mu, \qquad (1.65)$$

as is the coupling to photons,

$$= -i\frac{2}{3}e\gamma^\mu. \qquad (1.66)$$

Due to the $V - A$ structure of the charged weak currents, only the left-handed top couples to the $W^\pm$, with coupling

$$= ig\gamma^\mu(1 - \gamma^5)V_{tb}, \qquad (1.67)$$

where the value of V_{tb} is given in Eq. (1.57). The top couples to the Z with unequal left

and right-handed components, given by

$$= \frac{ig}{2\cos\theta_W}\gamma^\mu(v_t - a_t\gamma^5), \quad (1.68)$$

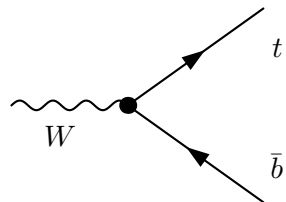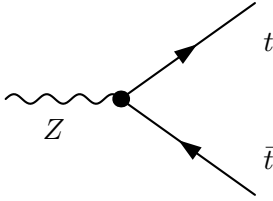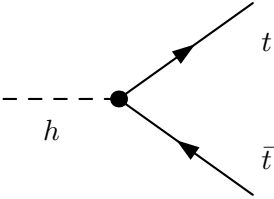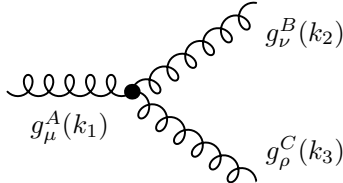where $v_t = T_{3t} - 2Q_t\sin^2\theta_W \simeq 0.19$ and $a_t = T_{3t} = 1/2$. Finally, it couples to the Higgs boson with a Yukawa-type interaction $\bar{\psi}\psi\phi$,

$$= y_t = \frac{\sqrt{2}m_t}{v}. \quad (1.69)$$

All of these couplings are flavour-conserving, with the exception of the charged-current interaction with the $W^\pm$. Since we are interested in top quark production at hadron colliders, the QCD triple gluon vertex will also be relevant for our discussion. Its Feynman rule is

$$= g_s f^{ABC}[g_{\mu\nu}(k_1-k_2)_\rho + g_{\nu\rho}(k_2-k_3)_\mu + g_{\rho\mu}(k_3-k_1)_\nu], \quad (1.70)$$

where all momenta are defined as towards the vertex.

The structure of the top quark couplings is identical to those of the other quarks, but the top enjoys properties unique amongst the quarks, namely its large coupling to the Higgs boson ($y_t \simeq 1$ in the SM) which suggests it plays a special role in electroweak symmetry breaking, and its large coupling to b-quarks (V_{tb} has been measured to be very close to 1), an observation which is unexplained in the SM. For these reasons and others, the top is often viewed as a possible window to physics beyond the Standard Model (indeed, this is the subject of this thesis). However, before turning our attention to BSM physics, we conclude this chapter with a discussion of the main production mechanisms for top quarks at hadron colliders.

1.4.1 Top pair production

By far the dominant production mechanism for top quarks in hadron collisions is top pair production $pp/p\bar{p} \to t\bar{t}$. The main contributions to this process come from QCD;
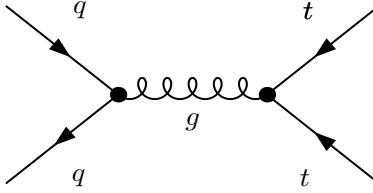
Figure 1.6: *The leading order Feynman diagram for $q\bar{q} \to t\bar{t}$ in the SM.*

production through intermediate Z bosons are negligible because the $t\bar{t}$ threshold is far from the Z pole, while QED contributions are parametrically suppressed by $(\alpha/\alpha_s)^2$. At leading-order in α_s, the partonic subprocesses $q\bar{q} \to t\bar{t}$ and $gg \to t\bar{t}$ both contribute. For the former, the partonic cross-section is, averaging (summing) over initial (final) state spins and colours:

$$\sigma_{q\bar{q}\to t\bar{t}} = \frac{g_s^4}{108\pi s}\beta(3-\beta^2),\tag{1.71}$$

where $\beta = \sqrt{1 - 4m_t^2/s}$ is the velocity of the top quark in the centre of mass frame (generically referred to as the 'threshold variable'). The leading-order Feynman$^{\|}$ diagram for the $q\bar{q}$ channel is sketched in Fig. 1.6. For the gg channel, we have:

$$\sigma_{gg\to t\bar{t}} = \frac{g_s^4}{768\pi s}\left(31\beta^3 - 59\beta + (33 - 59\beta + (33 - 18\beta^2 + \beta^4))\log\frac{1+\beta}{1-\beta}\right).\tag{1.72}$$

Feynman diagrams for this process are sketched in Fig. 1.7.

To obtain the full hadron-level cross-section, we convolute these expressions with the parton densities, as in Eq. (1.64). Displaying a closed-form expression for the hadron-level cross-section would thus require functional forms for the parton densities $f_g(x; Q^2)$ and $f_q(x; Q^2)$ used to fit the data. Here we simply discuss the numerical results, obtained from numerical tables of the pdf data.

The relative contributions of the partonic subprocesses are determined by the nature of the incoming hadrons. At the Tevatron $p\bar{p}$ collider, antiquarks exist as valence quarks in the initial state, so $q\bar{q} \to t\bar{t}$ is the dominant subprocess: it contributes around 85% of the total cross-section, the remainder is made from gluon-fusion. At a centre of mass energy of $\sqrt{s} = 1.96$ TeV, the leading order cross-section is calculated to be around 7 pb, for $\mu_R = \mu_F = m_t$ and using the CTEQ6l1 parton sets. At the LHC, antiquarks only appear as sea quarks, whilst the large kinematic reach means the proton is resolved down to much smaller momentum fraction $x_{min} \sim 10^{-5}$. In this regime the gluon luminosity becomes dominant, so the gg channel contributes up to 90% of the total cross-section. At a centre

---

$^{\|}$Our convention for Feynman diagrams is to represent the flow of charge through diagrams with arrows, in keeping with the Feynman-Stuckleberg interpretation of antimatter as matter under a time reversal transformation.
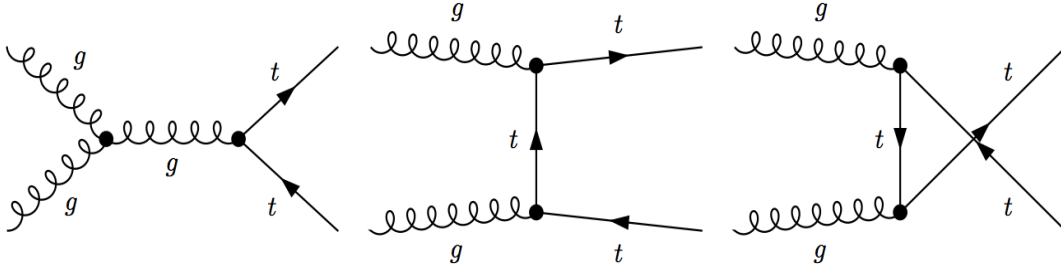
Figure 1.7: *The leading order Feynman diagrams for* $gg \to t\bar{t}$ *in the SM. Diagrams with these topologies are generically labelled (from left to right): s, t and u-channel diagrams.*

of mass energy of 7 TeV, the leading-order cross-section is around 100 pb [101].

Higher-order corrections

Understanding the effects of higher-order radiative corrections is necessary for obtaining precise Standard Model cross-section predictions. The size of (as yet) uncalculated higher-order effects can be estimated by noting the change in the cross-section with respect to scale variations. Leading-order estimates are typically correct within a factor of two, i.e. they provide a good ballpark estimate, but, owing to the fact that they include information about appropriate scale choices that should absorb the large logarithms that occur at higher orders, next-to-leading order (NLO) and often higher still corrections must be included for truly accurate estimates. They can be approximately included by defining a K-factor

$$K = \frac{\sigma(pp \to X)_{(N)NLO}}{\sigma(pp \to X)_{LO}}, \tag{1.73}$$

The higher-order estimate is then simply calculated by multiplying ('reweighting') the leading-order estimate by the K-factor. For top-pair production, the current 'state-of-the-art' SM prediction is the full next-to-next-to leading order estimate, which includes the resummation of terms involving soft gluon emissions to next-to-next-to-leading logarithmic accuracy (shorthand NNLO+NNLL) leading to the following values [102–104]:

$$\sigma(pp \to t\bar{t} + X) = 172.0 \,^{+4.4}_{-5.8} \text{ (scale)} \,^{+4.7}_{-4.8} \text{ (pdf)} \quad \sqrt{s} = 7 \text{ TeV}$$
$$\sigma(pp \to t\bar{t} + X) = 245.8 \,^{+6.2}_{-8.4} \text{ (scale)} \,^{+6.2}_{-6.4} \text{ (pdf)} \quad \sqrt{s} = 8 \text{ TeV} \tag{1.74}$$
$$\sigma(p\bar{p} \to t\bar{t} + X) = 7.164 \,^{+0.11}_{-0.20} \text{ (scale)} \,^{+0.17}_{-0.12} \text{ (pdf)} \quad \sqrt{s} = 1.96 \text{ TeV}$$

As well as the total cross-section, it is useful to study the dependence of the cross-section on kinematic observables that can be measured at colliders. The most commonly studied

variables are briefly outlined:

- The invariant mass, defined as

$$m^2 = \left(\sum_i E_i\right)^2 - \left(\sum_i \boldsymbol{p_i}\right)^2 \tag{1.75}$$

  where the sum is over all final state particles i. Final state invariant mass distributions are the classic way of searching for new particles. A peak in the $t\bar{t}$ invariant mass distribution at high $m_{t\bar{t}}$ would be an unambiguous signal of a new resonance decaying to top quarks.

- A related kinematic quantity is the transverse momentum p_T of the top; large-p_T events correspond to events in the high-energy region, where possible new physics effects are most likely to lie.

- The distribution of particles throughout the geometry of the detector is usually specified in terms of the rapidity y, defined as

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z}\right). \tag{1.76}$$

  This is typically used as a geometrical proxy for polar-angle θ^{**} as, unlike θ, it is additive under Lorentz boosts in the z-direction.

Top quark differential distributions have been calculated at NLO and are now fully automated in various Monte Carlo event generator programs [105–108]. Full phase-space results (at parton level) for top quark differential distributions are now available at NNLO QCD [109, 110], however they are not yet implemented in a Monte Carlo simulation such that they can be interfaced to a parton shower and implemented in a realistic experimental cutflow. To illustrate the importance of NLO corrections, in Fig. 1.8 we plot kinematic distributions in σ at LO and NLO. Uncertainties related to scales and pdfs have not been shown, the point is merely to illustrate that NLO corrections are large (nearing 50% in some bins) which highlights the need to include them.

1.4.2 Charge asymmetries

An important probe of the Standard Model in top pair production is through charge asymmetries [111–115]. The most well-known of these is the so-called 'forward-backward'

**Most collider experiments use a spherical co-ordinate system, where θ is the angle between the beam (z-axis) and the particle track, and ϕ is the azimuthal angle between the track and the vertical, i.e. looking down the beam.
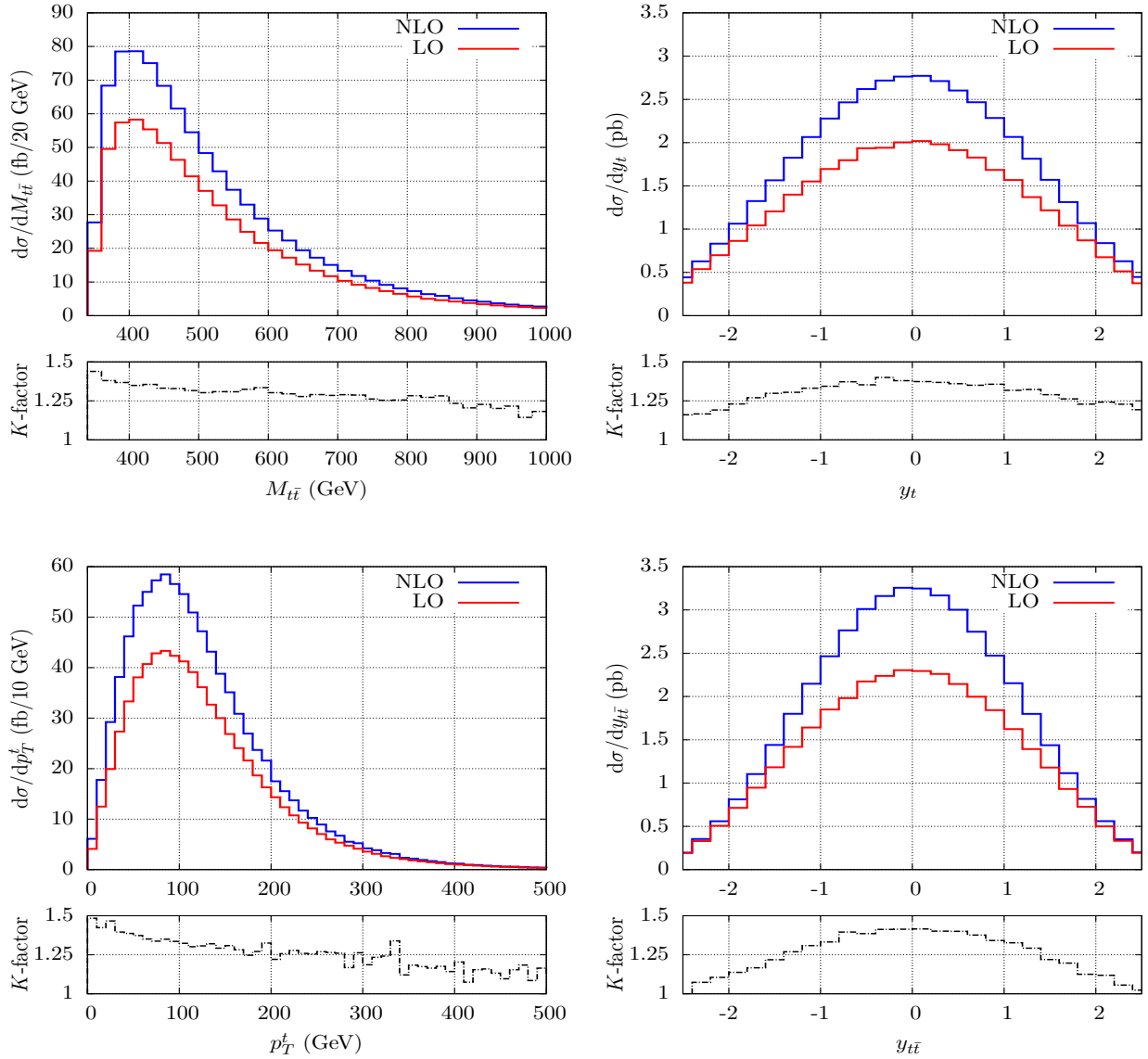
Figure 1.8: *Cross-section distributions in $pp \to t\bar{t}$ collisions at the LHC at NLO and LO, with associated bin-by-bin K-factors, as calculated with MCFM [107]. Here $y_{t\bar{t}} \equiv y_t - y_{\bar{t}}$.*

asymmetry in proton-antiproton collisions, which is most conveniently expressed as a difference between the number of top pairs in the forward direction (parallel with the incoming proton) and the backward direction (antiparallel with the incoming proton):

$$A_{FB} = \frac{N(\Delta y > 0) - N(\Delta y < 0)}{N(\Delta y > 0) + N(\Delta y < 0)}, \tag{1.77}$$

where $\Delta y = y_t - y_{\bar{t}}$. An asymmetry arises in the subprocess $q\bar{q} \to t\bar{t}$ due to terms which are odd under the interchange $t \leftrightarrow \bar{t}$ (with initial quarks fixed), specifically from the interference between the tree-level diagram for $q\bar{q} \to t\bar{t}$ and the 1-loop 'box' diagram, and interference between the real emission contributions for $q\bar{q} \to t\bar{t}g$. Thus, the asymmetry originates at next-to-leading order in QCD. The SM prediction at NNLO QCD is $A_{\text{FB}} =$

Figure 1.9: *Exaggerated schematic of the origin of the asymmetries A_{FB} at the Tevatron (left), and A_C at the LHC (right).*

7.24+1.04-0.67 [116], where the errors are from scale variation.

A different, but related, asymmetry can be defined at the LHC, where the charge symmetric initial state does not define a 'forward-backward' direction. Instead, a *central charge asymmetry* A_C can be defined

$$A_C = \frac{N(\Delta|y| > 0) - N(\Delta|y| < 0)}{N(\Delta|y| > 0) + N(\Delta|y| < 0)}, \tag{1.78}$$

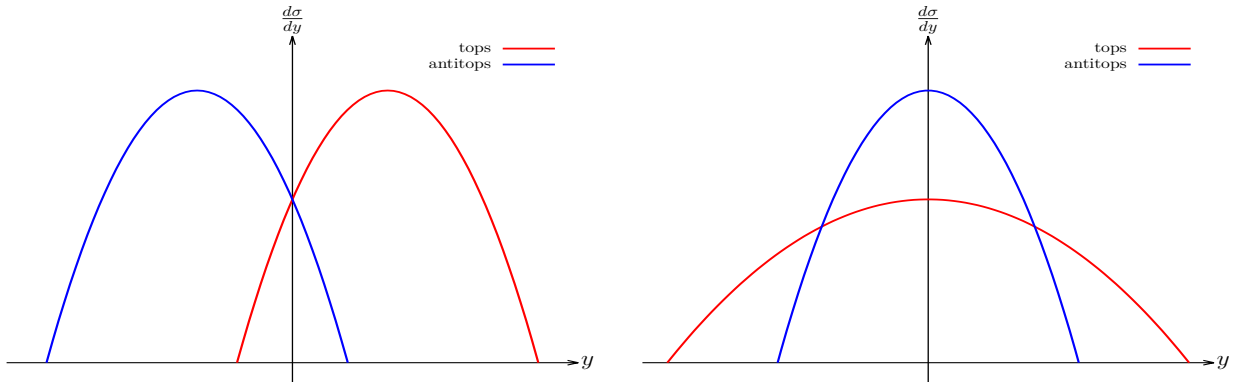where $\Delta|y| = |y_t| - |y_{\bar{t}}|$. This definition makes use of the fact that in $q\bar{q} \to t\bar{t}$ the quark in the initial state is almost always a valence quark and is likely to carry more longitudinal momentum than the antiquark, which is always a sea quark. The net result is that tops, being more correlated with the direction of the initial state quarks, tend to be produced at larger absolute rapidities than antitops. However, at LHC energies, gluons dominate the beam composition, so the $gg \to t\bar{t}$ channel, for which $A_C = 0$, dominates the cross-section. This means A_C is much more diluted than A_{FB}. Its SM prediction is $A_C = 0.0123 \pm 0.0005$ [111], which includes NLO QCD and electroweak corrections. The asymmetries at the Tevatron and the LHC are visualised in Fig. 1.9.

1.4.3 Single top production

The next-most-dominant way of producing top quarks at hadron colliders is the single-top process, which can be sub-categorised into the purely electroweak processes $q\bar{q}' \to t\bar{b}$ and $qb \to tq'$, mediated by W bosons in the s [117–119] and t-channel [120–124], and the electroweak+QCD process $gb \to tW$; referred to as Wt-associated production [125–131]. Feynman diagrams for both cases are shown in Figs. 1.10 and 1.11.

The s-channel cross-section has a relatively large rate at the Tevatron, but at the LHC it is much rarer than its t-channel counterpart, because it is initiated by antiquarks and
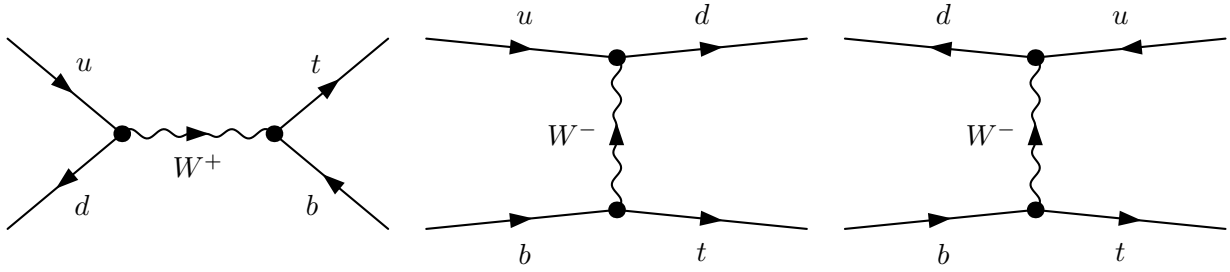
Figure 1.10: *The leading order Feynman diagrams for electroweak single top production in the SM. The corresponding antitop diagrams are constructed by reversing the fermion arrows.*

so is suppressed by the initial parton densities. The signature for s-channel production is a pair of b-tagged quarks, one originating from the primary vertex and one from the decay of the top quark, a high p_T lepton, and missing transverse energy, corresponding to a neutrino from the leptonic top quark decay. It remains a challenging channel to reconstruct, however, due to its small event rate and large backgrounds, namely from top pair and W+jets. The leading order partonic cross-section for s-channel top production is

$$\sigma_{u\bar{d}\to t\bar{b}} = \frac{|V_{ud}|^2|V_{tb}|^2 g^4 (s - m_t^2)^2 (2s + m_t^2)}{384\pi s^2 (s - M_W^2)^2}. \tag{1.79}$$

In t-channel production, in order to produce a top quark, the spacelike W must be highly off-shell, and so there is a large momentum transfer between the outgoing partons, hence the light quark tends to recoil against the heavy top, leading to an untagged jet in the forward region of the detector. Moreover, the exchange of a color singlet between the two outgoing partons means there is relatively little QCD radiation in the region between them, leading to suppressed central jet activity between the top quark decay products and the jet from the light quark, known as a *rapidity gap*. Though this defines a very clear experimental signature, at the theoretical level there exists some ambiguity in the parton-level definition of this process. One may choose to define the incoming b-quark as originating directly from the incoming proton, using a so-called 5-flavour scheme for the proton pdf, leading to the $2 \to 2$ topology as shown in Fig. 1.10. Alternatively, one may treat the b-quark as the product of the collinear splitting of a gluon ($g \to b\bar{b}$) in the initial state, leading to a $2 \to 3$ event topology.

Formally, these two treatments should lead to the same cross-section prediction, but differ when truncated at fixed-order in perturbation theory, in particular due to the accuracy at which the logarithms originating from the gluon splitting are resummed, and the treatment of these splittings in the evolution of the pdfs. The leading-order parton level cross-section for t-channel production in the 5-flavour scheme, in both the $ub \to dt$ and
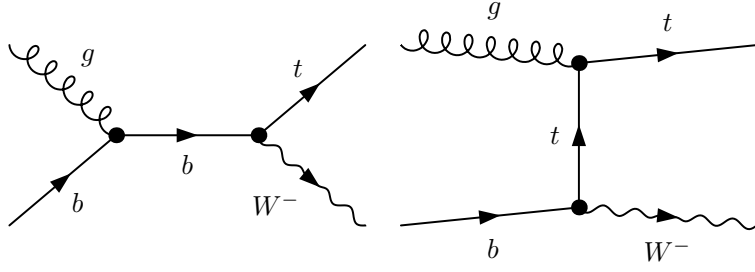
Figure 1.11: *The leading order Feynman diagrams for Wt-associated production in the SM. The corresponding antitop diagrams are constructed by reversing the fermion arrows.*

$d\bar{b} \to \bar{u}t$ channels are:

$$\sigma_{ub \to dt} = \frac{|V_{ud}|^2 |V_{tb}|^2 g^4 (s - m_t^2)^2}{64\pi s M_W^2 (s - m_t^2 + M_W^2)}$$

$$\sigma_{d\bar{b} \to \bar{u}t} = \frac{|V_{ud}|^2 |V_{tb}|^2 g^4 \left((s - m_t^2)^2 (2s + m_t^2) - M_W^2 (2s + 2M_W^2 - m_t^2) \log \frac{s + M_W^2 - m_t^2}{M_W^2} \right)}{4\pi s^2 M_W^2}.$$

$$(1.80)$$

Finally, for Wt-associated production, the cross-section in the 5-flavour scheme takes the form

$$\sigma_{gb \to Wt} = \frac{|V_{ud}|^2 |V_{tb}|^2 g^2 g_s^2}{384 s^3 M_W^2} \Bigg(-3((m_t^2 - 2M_W^2)s + 7(m_t^2 - M_W^2)(m_t^2 + 2M_W^2))\lambda^{1/2}(s, m_t^2, M_W^2)$$

$$+ 2(m_t^2 + 2M_W^2)(s^2 + 2(m_t^2 - M_W^2) + 2(m_t^2 - M_W^2)) \log \left(\frac{s + m_t^2 + M_W^2 + \lambda^{1/2}}{s + m_t^2 + M_W^2 - \lambda^{1/2}} \right) \Bigg)$$

$$(1.81)$$

where $\lambda(x, y, z) = x^2 + y^2 + z^2 - 2xy - 2xz - 2yz$ is the Källén function [132].

The pure electroweak single top production processes typically have cross-sections an order of magnitude smaller than for top pair production. Although the available phase space for producing one top instead of two is much larger, the matrix elements are parametrically suppressed by the strength of the electroweak coupling relative to the strong coupling. For the same reason, the cross-sections are more stable against higher-order corrections, and K-factors for s and t-channel production are more flat in differential distributions and scale choices, typically at the 10-20% level. The most up-to-date calculations for electroweak single-top production are at approximate NNLO, although there are different definitions of this term. One calculation calculates in the so-called structure function approximation, where only factorisable vertex corrections are considered [124]. The remaining terms are colour suppressed $\sim 1/N_c^2$ and kinematically subdominant. Another approach is to expand the resummed leading-order cross-section to $\mathcal{O}(\alpha_s^2)$ [120]. Both calculations are in general
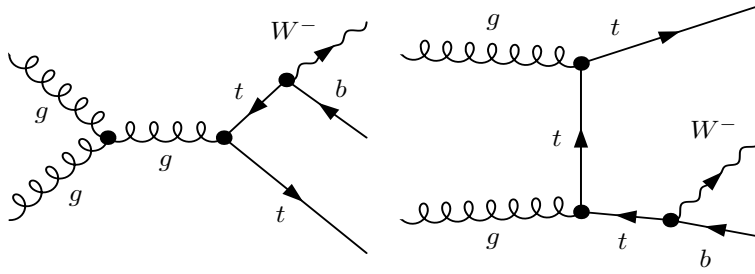
Figure 1.12: *Real emission contributions to Wt-associated production at NLO. These also correspond to diagrams for $t\bar{t}$ production, where the antitop has decayed.*

agreement. The latter yields for the t-channel:

$$\sigma(pp \to tq + X) = 65.7 \; ^{+1.9}_{-1.9} \quad \sqrt{s} = 7 \text{ TeV}$$
$$\sigma(pp \to tq + X) = 87.1 \; ^{+0.24}_{-0.24} \quad \sqrt{s} = 8 \text{ TeV} \tag{1.82}$$
$$\sigma(p\bar{p} \to tq + X) = 2.06 \; ^{+0.13}_{-0.13} \quad \sqrt{s} = 1.96 \text{ TeV}.$$

where the uncertainties quoted have added scale and pdfs in quadrature, and for the s-channel:

$$\sigma(pp \to tb + X) = 4.5 \; ^{+0.2}_{-0.2} \quad \sqrt{s} = 7 \text{ TeV}$$
$$\sigma(pp \to tb + X) = 5.5 \; ^{+0.2}_{-0.2} \quad \sqrt{s} = 8 \text{ TeV} \tag{1.83}$$
$$\sigma(p\bar{p} \to tb + X) = 1.03 \; ^{+0.05}_{-0.05} \quad \sqrt{s} = 1.96 \text{ TeV}.$$

These values are summed over the top and antitop channels. At the Tevatron, owing to its charge symmetric initial states, both channels contribute equally, while at the LHC the relative top/antitop contributions are 65% to 35% for t-channel, and 69% to 31% for s-channel.

Since the Wt process is QCD initiated, it is expected to receive sizeable corrections from higher-order terms. However, an ambiguity arises when one tries to define an NLO estimate for this process. Generically, NLO corrections result from both virtual 'loop' corrections, and emission of real particles. The latter type in Wt production include diagrams of the form shown in Fig. 1.12: which are also present in resonant top-pair production, with one top quark decay $t \to Wb$. When this intermediate top quark goes on-shell, the Wt cross-section becomes of the order of the $t\bar{t}$ one, which is an order of magnitude larger. In this regime, such a large K-factor means a perturbative definition of the Wt process is ill-defined. The question is then, do such diagrams belong to Wt or $t\bar{t}$, and can one make an NLO definition of Wt production that avoids the interference with resonant top-pair production? This problem has been studied in some detail [128, 133–137].

Two well-known prescriptions for removing the effects of top pair production in Wt at

NLO are *diagram removal* and *diagram subtraction*. The former removes the diagrams of the form of Fig. 1.12 at the amplitude level, so they do not enter the calculation. The latter subtracts their contributions from the final cross-section. The difference between the NLO Wt cross-section predictions from these two methods thus provides a measure of the interference effect of the diagrams of Fig. 1.12 [129, 133]. Reasonable choices of experimental cuts can be made to minimise this interference (an obvious choice, for instance, would involve an invariant mass cut close to the $t\bar{t}$ threshold), so that Wt and $t\bar{t}$ can, for all practical purposes, be considered as separate processes[††].

However one decides to define the Wt-associated production process, the cross-section is too small at the Tevatron to be of any phenomenological relevance. At the LHC, however, the approximate NNLO cross-section, defined from the NNLO expansion of the NNLL-resummed cross-section, is:

$$
\begin{aligned}
\sigma(pp \to Wt + X) &= 15.5 \ ^{+1.2}_{-1.2} \quad \sqrt{s} = 7 \text{ TeV} \\
\sigma(pp \to Wt + X) &= 22.1 \ ^{+1.2}_{-1.2} \quad \sqrt{s} = 8 \text{ TeV}.
\end{aligned}
\tag{1.84}
$$

Since the process is initiated by a gluon and a b-quark, and $f_b(x) = f_{\bar{b}}(x)$ in the proton, the top/antitop contributions are equal.

1.4.4 Higher-order processes

As the LHC probes kinematic regions inaccessible to previous colliders such as the Tevatron, new event topologies with a higher multiplicity of hard partons become increasingly commonplace. Of special interest for top physics are processes where a top quark pair is produced in association with an additional particle in the hard process, dubbed *higher-order* because they already have a $2 \to 3$ topology at tree-level. These processes probe directly the top couplings of Eqs. (1.65)-(1.69), allowing for a model-independent way of constraining new top interactions. For instance, a measurement of top pairs in association with a Z boson directly probes the $t\bar{t}Z$ [138, 139] coupling, allowing contact to be made with precision LEP observables. In principle this could be extracted from the simple $pp \to Z \to t\bar{t}$ process, but this signal is drowned out by the much larger QCD $t\bar{t}$ rate.

Top pairs produced in association with Higgs bosons ($t\bar{t}H$) [140, 141] are of particular interest in this regard, because they allow for a model-independent extraction of the top quark Yukawa coupling, thus offering discriminating power between models where much of the top quark mass is generated from a non-SM mechanism. These processes have small rates, typically $\mathcal{O}(100 \text{ fb})$, and are thus experimentally challenging. Nonetheless, significant progress has been made towards their discovery in LHC Run I [142, 143], and

---

[††]For an alternative viewpoint, see e.g. Ref. [136].

the high statistics forecast for the LHC lifetime suggest they can ultimately be measured with similar precision to the leading order processes discussed above.

1.4.5 Top quark decay

The unique properties of the top quark stem largely from the characteristics of its decay. The top quark is the only quark in the Standard Model with a mass larger than that of the W boson. Hence it can decay directly through the process $t \to Wb$. The mass difference $m_t - M_W \sim 90$ GeV means that the allowed phase-space for the decay is large, and so the top quark decays before the strong interaction can bind it into hadrons. At next-to-leading order in α_s, the top decay width is given by [144]

$$\Gamma_t = \frac{G_F m_t^3}{8\pi\sqrt{2}} \left(1 - \frac{M_W^2}{m_t^2}\right) \left(1 + 2\frac{M_W^2}{m_t^2}\right) \left[1 - \frac{2\alpha_s}{3\pi}\left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right]. \tag{1.85}$$

where terms of order $m_b^2/m_t^2, \alpha_s^2$ and $(\alpha_s/\pi)M_W^2/m_t^2$ have been neglected, and it is assumed $|V_{tb}|^2 = 1$. For $m_t = 173.3$ GeV this gives a value of $\Gamma_t \sim 1.3$ GeV, corresponding to a top quark lifetime of 0.5×10^{-24} s. Despite the large phase space available for the decay, the top still satisfies the narrow width approximation $\Gamma \ll m$. This means that one can make the replacement of the propagator

$$\frac{1}{((s - m^2)^2 + (m\Gamma)^2)} \to \frac{\pi}{m\Gamma}\delta(s - m^2), \tag{1.86}$$

in the squared matrix element. This substantially simplifies the calculation of decay amplitudes. Nonetheless, the narrow width approximation should be treated with care, as it is not valid for observables whose main contributions originate from regions of phase space where the top is far off-shell. In addition, the raw top width is a difficult quantity to measure at hadron colliders, without making assumptions (such as a SM-like cross-section). Other observables relating to top decay can be much more precisely measured.

For instance, the fraction of events in which the top decays to W-bosons with a given helicity: left-handed, right-handed or zero-helicity, can be expressed in terms of helicity fractions, which for leading order with a finite b-quark mass are

$$
\begin{aligned}
F_0 &= \frac{(1 - y^2)^2 - x^2(1 + y^2)}{(1 - y^2)^2 + x^2(1 - 2x^2 + y^2)} \\
F_L &= \frac{x^2(1 - x^2 + y^2) + \sqrt{\lambda}}{(1 - y^2)^2 + x^2(1 - 2x^2 + y^2)} \\
F_R &= \frac{x^2(1 - x^2 + y^2) - \sqrt{\lambda}}{(1 - y^2)^2 + x^2(1 - 2x^2 + y^2)},
\end{aligned}
\tag{1.87}
$$

where $x = M_W/m_t$, $y = m_b/m_t$ and $\lambda = 1 + x^4 + y^4 - 2x^2 y^2 - 2x^2 - 2y^2$. A desirable feature of these quantities is that they are relatively stable against higher order corrections, so the associated scale uncertainties are small. The Standard Model NNLO estimates for these are: $\{F_0, F_L, F_R\} = \{0.687 \pm 0.005, 0.311 \pm 0.005, 0.0017 \pm 0.0001\}$ [145, 146], i.e. the uncertainties are at the per mille level. The large enhancement of decays to longitudinal W bosons results from the Goldstone boson equivalence theorem, which states that in the limit $s \gg M_W$, W scattering is dominated by the longitudinal components, so Ws may be approximated by Goldstone scalars in calculations.

Another unique feature of the top quark is that its decay width is much larger than the QCD spin decorrelation width $\Lambda_{QCD}^2/m_t \sim 0.1$ MeV, therefore the correlation between the spins of tops and their decay products is completely preserved, and spin correlations in, for example, top pair production can be measured directly through spins of the decay leptons (selecting dilepton events).

1.5 Summary

To summarise, the top quark plays a special role among the Standard Model fermions, and, owing to its unique experimental properties, offers a valuable hadron collider testing ground for many SM predictions. However, the real interest in the top quark stems from its role in potential TeV scale new physics, perhaps within reach of the LHC. This is the subject of the next chapter.

2 The top quark beyond the Standard Model

2.1 Introduction

Despite the vast list of experimentally verified predictions of the SM over the last forty years, it still paints a somewhat unsatisfying picture of Nature. In order for it to be predictive, it requires fixing the values of 18* arbitrary parameters from experiment. These parameters span several orders of magnitude, with no apparent pattern between them. Any fundamental theory worth its salt ought to be able to predict the values of these numbers, or at least relate them in terms of a smaller subset of more fundamental parameters.

Beside the aesthetic issue of the large number of free parameters, the Standard Model also has some deep structural problems that have motivated new physics model building for the last few decades. Owing to its unique properties among the SM fermions, the top quark has played a special role in most of these scenarios. Indeed, the potential for using measurements of top quark couplings to place bounds on the effects of new physics is the main topic of this thesis.

This chapter is structured as follows: In section 2.2 I will outline the main motivations for physics beyond the Standard Model (BSM). In section 2.3 I discuss a few of the most popular scenarios of BSM physics, and the role the top quark plays in them. In section 2.4 I discuss some generalities of using effective field theories to parameterise the effects of heavy degrees of freedom on low energy observables, before moving onto discussing the formulation of the Standard Model as an effective theory in section 2.5, and the parts of that EFT that are relevant for top quark physics in section 2.6. Conclusions are presented in section 2.7.

2.2 Motivations for physics beyond the Standard Model

2.2.1 The hierarchy problem

Perhaps the best-known motivation for physics beyond the SM is the hierarchy problem: the vast difference between the electroweak scale and the Planck scale where quantum gravity becomes important: $v/M_{Pl} \sim 10^{-16}$. This large mass hierarchy is not a specific problem of the Glashow-Salam-Weinberg model, but a general feature of theories containing fundamental scalars. To see this, we return to the simple case of a real scalar field in four dimensions.

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - V(\phi) \quad \text{where} \quad V(\phi) = \frac{1}{2}\mu^2 \phi^2 + \frac{1}{4}\lambda \phi^4 \tag{2.1}$$

---

*Excluding the QCD θ-parameter, which will be discussed later in this chapter.
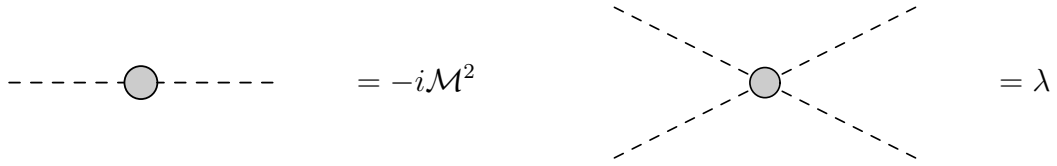
Figure 2.1: *Feynman rules for the 1-loop mass and coupling constants defined in Eqs. (2.4)-(2.5).*

At tree-level the potential will simply correspond to the classical potential $V_{\text{tree}} = V(\phi)$. However, radiative corrections will modify this relation, and the *true* potential; that is, the one that the vacuum expectation value seeks to minimise, is the so-called *effective potential* V_{eff}, which will receive radiative corrections.

The radiative corrections originate from the effects of virtual particle emission and absorption on the interaction energy, so in principle includes all one-particle irreducible diagrams with any number of external legs $n > 2$[†]. The potential involves only non-derivative terms in the Lagrangian, so the momenta of the external legs can be taken to be zero without loss of generality, i.e. calculating the radiative corrections to the scalar potential amounts to summing up all 1PI diagrams with zero external momentum (for a more formal proof of this statement, see e.g. Refs [148, 149]). This can be done order-by-order in perturbation theory. To isolate the quantum corrections, it is useful to split ϕ into a classical 'background' or external field ϕ_c, corresponding to the field in the tree potential, and quantum corrections $\delta\phi$.

$$\phi = \phi_c + \delta\phi. \tag{2.2}$$

Since we will encounter divergent loop momenta, we must define renormalisation conditions for the couplings of the theory to absorb them. Conventionally, one defines the renormalised mass of the scalar field ϕ as the 'pole' mass, i.e. the negative of the inverse propagator at zero momenta

$$\mathcal{M}^2 \equiv -\frac{\delta^2 \mathcal{L}}{\delta\phi_c^2}\bigg|_{\phi_c=0}, \tag{2.3}$$

while the renormalised coupling is defined from the 4-point function at zero momentum:

$$\lambda \equiv -\frac{\delta^4 \mathcal{L}}{\delta\phi_c^4}\bigg|_{\phi_c=0}. \tag{2.4}$$

In the spontaneously broken theory, the subtraction point is shifted from $\phi_c = 0$ to $\phi_c = \langle\phi\rangle$

[†]Vacuum energy graphs with $n = 0$ just shift the energy by a constant, diagrams with $n = 1$ can be absorbed into a shift of the field [147].

so that

$$\mathcal{M}^2 \equiv -\frac{\delta^2 \mathcal{L}}{\delta \phi_c^2}\bigg|_{\phi_c = \langle \phi \rangle} = \mu^2 + \frac{\lambda}{2} \langle \phi \rangle^2 . \tag{2.5}$$

This means that the renormalised masses and couplings will in general be functions of $\langle \phi \rangle$. The corresponding Feynman rules are shown in Fig. 2.1.

As mentioned above, the all-order effective potential is given by the sum of all 1PI vacuum diagrams with zero external momentum. At one-loop then, calculating the contributions to V_{eff} amounts to summing up the vacuum 'bubble' diagrams of the form shown below [149],

$$V_{\text{eff}}^{\text{1-loop}}(\phi_c) =$$
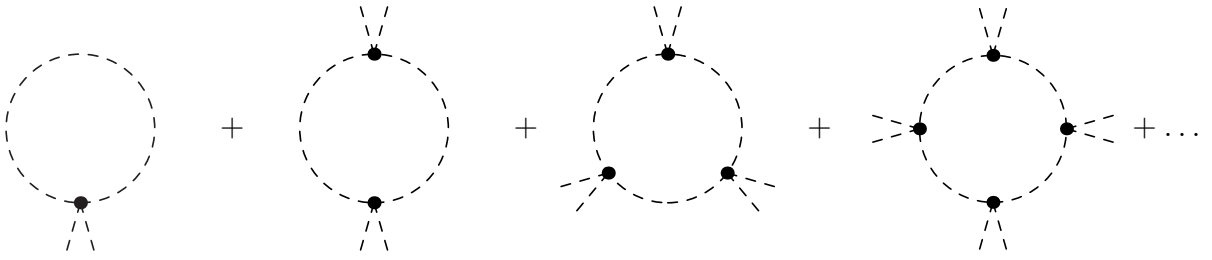
leading to a geometric series which can be resummed, giving (up to constant terms) the well-known one-loop Coleman-Weinberg effective potential [148]:

$$V(\phi_c) = V_{\text{tree}} + V^{(1)}(\phi_c)$$
$$V^{(1)}(\phi_c) = \frac{\Lambda^2}{32\pi^2}\mu^2 + \frac{\lambda\phi_c^2\Lambda^2}{32\pi^2} + \frac{(\mu^2 + \lambda\phi_c^2)^2}{64\pi^2}\left[\log\frac{(\mu^2 + \lambda\phi_c^2)}{\Lambda^2} - \frac{1}{2}\right]. \tag{2.6}$$

Clearly, this expression is divergent: the mass term has a divergence proportional to the UV cutoff Λ^2, and the quartic coupling has a logarithmic divergence. These may be absorbed into counterterms δ_μ and δ_λ specified by the renormalisation conditions of Eqs. (2.4) and (2.5), so that the full potential at 1-loop order is then

$$V(\phi_c) = \mu_{\text{renorm}}^2 \phi_c^2 + \frac{\lambda_{\text{renorm}}}{4}\phi_c^4 = V_{\text{tree}} + V^{(1)}(\phi_c) + V_{\text{c.t.}} \tag{2.7}$$

This, however, means that the renormalized (physical) mass will receive corrections of the form

$$\mu_{\text{renorm}}^2 = \mu_0^2 + \frac{\lambda_0\Lambda^2}{32\pi^2} + \frac{\lambda_0\mu_0^2}{64\pi^2}\left[\log\left(\frac{\mu_0^2 + \frac{\lambda\langle\phi\rangle^2}{2}}{\Lambda^2}\right)^2 - \frac{1}{2}\right] - \delta_\mu. \tag{2.8}$$

So in order to keep the renormalised mass of the same order as the bare mass $m_0 = \sqrt{-2\mu^2}$, we require a cancellation between the quadratically divergent term Λ and the

counterterm δ_μ (the logarithmic term remains of the same order since it is multiplied by μ_0^2). If the physical mass is to be much smaller than the cutoff Λ, we must assume a miraculous cancellation between contributions below the cutoff and the unknown UV degrees of freedom, parameterised by the counterterms, above the cutoff [150, 151].

This analysis can be applied to the μ^2 term in the Higgs potential of the Standard Model. Here the cutoff Λ denotes the generic scale at which the Standard Model is no longer valid, it could for instance represent the mass of a new heavy scalar. Since it can be subtracted off in mass renormalisation, it should not affect low-energy physics. To ensure this, however, requires an extraordinarily precise fine-tuning of parameters. Suppose, for instance, that the SM were valid all the way up to the Planck scale. To keep the renormalized mass at the 100 GeV scale one would need to arrange for the cancellation between the 'bare' mass, describing the low-energy theory, and the counterterms, describing unknown high-energy degrees of freedom. This cancellation would have to be precise to 16 orders of magnitude, and hold through several orders in perturbation theory. Even if one started without a tree level mass, and generated it radiatively, i.e. by just considering the quartic term in the tree-level potential, the large mass corrections would still be present, because the scalar mass renormalisation is additive, not multiplicative[‡].

We know of no other situation in physics where degrees of freedom separated by so many orders of magnitude would conspire to produce the phenomena that we observe. To calculate the Bragg diffraction angles on a crystal, for instance, one does not need to know the mass of the Z boson. In keeping with this separation of scales principle, it seems that the natural mass for a fundamental scalar in a theory is close to the cutoff of that theory. What mechanism is it, then, that keeps the Higgs so light? This is the hierarchy problem, and has been the main driving force for physics beyond the Standard Model for the last forty years.

One could restate the argument in a different way. The hierarchy problem is *not* the fact that there is a large difference between the electroweak scale and the Planck scale. One does not complain, for instance, about the large hierarchy ($\sim 10^6$) between the electron mass and the electroweak scale. This is because the electron mass term in the Standard Model originates from its chiral Yukawa coupling to the Higgs field:

$$\bar{L}\varphi e_R \rightarrow y_e(\nu_L, e_L)H e_R. \tag{2.9}$$

This is the only term in the SM Lagrangian that breaks the electron's chiral symmetry. Consequently, any radiative corrections that break chiral symmetry can only be propor-

---

[‡]Although other regularisation schemes such as dimensional regularisation 'hide' the UV divergences by lacking a UV cutoff Λ, this does not mean the hierarchy problem is an artefact of using a momentum-dependent regulator: it is a statement that parameters at one scale are sensitive to parameters at a vastly different scale, whether one expresses this scale in terms of a cutoff or not.

tional to positive powers of y_e. Setting $y_e = 0$ thus enlarges the symmetry group of the Standard Model. Approximate symmetries like this have physical consequences such as (approximately) conserved currents, so there is a 'natural' reason for m_e to take such a small value in relation to other relevant scales. y_e is an example of a 'technically natural' parameter [152]. The Higgs mass term $\mu^2(H^\dagger H)$, on the other hand, is not technically natural; since it is already invariant under any chiral transformation $H \to e^{i\theta\gamma_5} H$, so it has no natural reason for being so much smaller than naïve power counting would suggest.

2.2.2 Vacuum stability

Arguments for Naturalness of the Higgs mass as evidence for the need for new physics are convincing, but not incontrovertible. The Standard Model is a renormalisable field theory, which means that it is in principle a valid description of Nature from the electroweak scale all the way up to the Planck scale, where the degrees of freedom of quantum gravity will become important. It is thus possible, that there is no new physics in the region in between. Extrapolating the Standard Model across this many orders of magnitude, however, leads to an interesting implication for cosmology. To show this, we consider the 1-loop renormalisation group equations for the following Standard Model parameters, with $n_f = 6$ flavours of quark:

- The hypercharge coupling g':

$$\mu \frac{dg'}{d\mu} = \frac{41}{6} \frac{g'^3}{16\pi^2} \tag{2.10}$$

- The SU(2) gauge coupling g:

$$\mu \frac{dg}{d\mu} = -\frac{19}{6} \frac{g^3}{16\pi^2} \tag{2.11}$$

- The strong coupling constant g_s:

$$\mu \frac{dg_s}{d\mu} = -7 \frac{g_s^3}{16\pi^2} \tag{2.12}$$

- The Higgs quartic coupling λ:

$$\mu \frac{d\lambda}{d\mu} = \frac{1}{16\pi^2} \left(\frac{3g'^4}{8} + \frac{3g^2 g'^2}{4} + \frac{9g^4}{8} - 6y_t^4 - \lambda(3g'^2 + 9g^2 - 12y_t^2) + 24\lambda^2 \right) \tag{2.13}$$

- The top quark Yukawa coupling y_t:

$$\mu \frac{dy_t}{d\mu} = \frac{y_t}{16\pi^2} \left(\frac{9y_t^2}{2} - \frac{17g'^2}{12} - \frac{9g^2}{2} - 8g_s^3 \right). \tag{2.14}$$
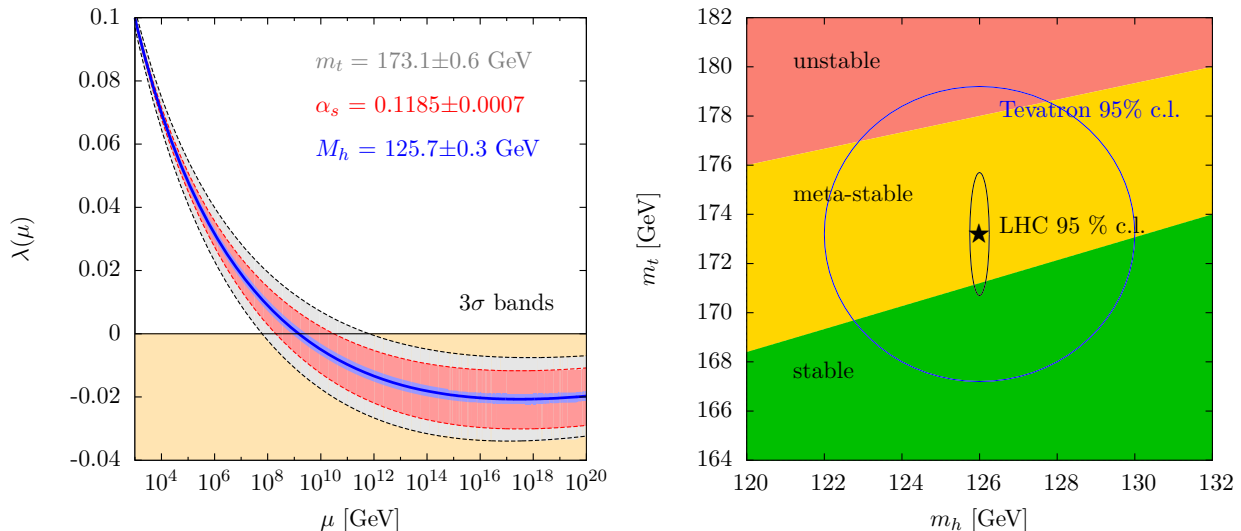
Figure 2.2: *Left: One-loop renormalisation group evolution of the SM Higgs self-coupling λ. Also shown are the 3σ bands for the three most dominant sources of uncertainty: the measured value of the top quark mass m_t, the strong coupling constant α_s, and the Higgs mass M_h. Right: Corresponding regions of stability, meta-stability and instability (as described in the text) of the electroweak vacuum, as a function of the Higgs and top quark masses, overlaid with the most recent best fit contours for both from the Tevatron and LHC.*

The evolution of the Higgs self-coupling is clearly most sensitive to the top quark Yukawa (all of its other fermionic couplings can be safely neglected). It is also sensitive to the large running of the strong coupling constant (indirectly through y_t) and to itself, λ. Substituting in as boundary conditions the values of SM couplings at the electroweak scale $\mu = v$, taken from Tab 1.1, and using $\lambda = m_h^2/2v^2$, we can straightforwardly solve for $\lambda(\mu)$. The running of the Higgs self-coupling is plotted on the left of Fig. 2.2.

The y_t^4 term tends to drive the Higgs potential negative at large renormalisation scales. The precise scale at which this happens is extremely sensitive to the values of the electroweak scale inputs, but it is clear that $\lambda < 0$ at some scale $\mu = \Lambda < M_{Planck}$. It is unclear how this should be interpreted. As mentioned in chapter 1, the absolute stability of the Higgs potential requires that $\lambda > 0$, otherwise the potential will be unbounded from below. In the SM alone, there are no terms which can rescue the boundedness of the potential: it is negative definite for $\mu > \Lambda$. However, one can reasonably assume that couplings between the Higgs sector and Planck scale physics, which would manifest in the Higgs potential as higher-dimensional operators such as $(\varphi^\dagger\varphi)^3$, will restore the boundedness of the potential. Still, this means that the electroweak vacuum is not a *true vacuum*: there is another vacuum at a much higher scale (perhaps at M_{Planck}) for which it is energetically favourable for the Universe to tunnel into. Based on whatever numerical value

the tunnelling rate ξ takes[§] , one can draw three possible consequences for the fate of the electroweak vacuum [154, 155]:

- $\xi = 0$:
  The tunnelling rate is exactly zero, and v is actually the true minimum of the Higgs potential. This means the Higgs potential is absolutely stable. It is apparently disfavoured by data, unless the top mass and Higgs mass are respectively somewhat smaller and larger than their current measurements suggest.

- $\xi < 1/T_{Universe}$:
  The tunnelling rate is non-zero, but with a decay lifetime larger than the current age of the Universe, which would explain why the tunnelling has not yet taken place.

- $\xi > 1/T_{Universe}$:
  The tunnelling rate is faster than the Hubble rate, meaning the Universe should have undergone a phase transition from the electroweak vacuum to the true one, sometime between the Big Bang and today.

One can express these different possibilities in terms of m_t and M_h, as shown on the right of Fig. 2.2. The current measurements place us squarely in the metastable region. This fact alone is unremarkable (although if we were in the unstable region we ought to have a good explanation). However, it is a peculiar outcome for the evolution of the Universe. Shortly after the Big Bang, the Universe was in a state of very large free thermal energy. At some point the Universe must have then cooled enough to undergo a phase transition to the electroweak vacuum. It is then unknown what caused the Universe to choose the less energetically favourable of the two vacua, or what stabilised this false vacuum against quantum tunnelling and thermal fluctuations, which would have been much more likely in the early Universe (i.e. when ξ was much larger than $1/T_{Universe}$)[¶].

If one believes in Naturalness as a guiding principle, one should take this problem as seriously as the hierarchy problem. This apparent paradox may be taken as indirect evidence that new physics lies between v and M_{Planck} which stabilises the potential, because the preceding argument was only valid if there was no such new physics. At any rate, it shows that despite satisfying the requirement of renormalisability, the Standard Model alone paints a rather unsatisfactory picture of physics between v and M_{Planck}.

[§]An expression for the tunnelling/vacuum decay rate in terms of $\lambda(\mu)$ can be found in e.g. Ref. [153].
[¶]The state-of-the-art calculation of vacuum stability in the SM is at NNLO [156–159], rather than the leading-order one presented here, and a more comprehensive discussion of the issue can be found in those papers.
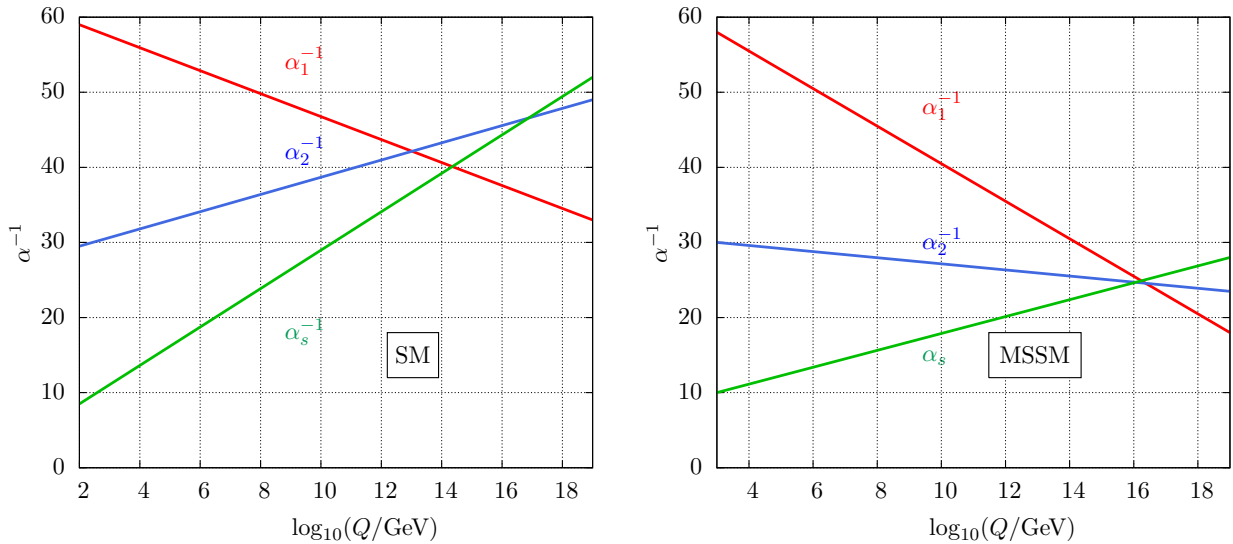
Figure 2.3: *Left: One-loop renormalisation group evolution of the SM gauge couplings. Right: Their renormalisation group flow in the Minimal Supersymmetric extension of the Standard Model (MSSM), showing unification at a scale $Q \sim 10^{16}$ GeV.*

2.2.3 Gauge coupling unification

The extraordinary accuracy of the electroweak theory in describing physics up to v has led to speculation that there may be an even larger unification scenario, in which the electroweak and strong interactions are unified under a single gauge group, with new degrees of freedom occupying that unification scale. Consider instead Eqs. (2.10)-(2.12). Although the coupling constants have very different values at the electroweak scale, they run in different directions, so perhaps unify at a single energy scale. The one-loop renormalisation group flow of g', g and g_s in the Standard Model is shown on the left of Fig. 2.3.

It can be seen that the couplings do not meet, and so no unification takes place. On the other hand, adding in new degrees of freedom between v and M_{Planck} will also alter the running of the couplings. One of the most well-studied candidates for new physics is surely supersymmetry. In the minimal supersymmetric extension of the MSSM, where every SM particle is supplemented with a superpartner of opposite spin (i.e. there is only one copy of the supersymmetry algebra), the couplings unify$^{\|}$ at a scale of around $Q \sim 10^{16}$ GeV [161–163]. This is a striking result, and is often considered one of the main motivations for supersymmetry as a candidate for new physics at the TeV scale.

$^{\|}$In fact, the couplings do not exactly unify in the MSSM, but exact unification can be achieved by adding appropriate threshold corrections at the unification scale, see e.g. Ref. [160].

2.2.4 Strong CP violation

Although there were 18 parameters in the Standard Model listed in the previous chapter, in principle one could write down a 19th ($\mathcal{CP}$-odd) term that could lead to experimental effects in the strong interaction. The fact that these effects have not been observed is known as the *strong CP problem*[**]. The origin of this problem lies in the fact that one can write down another gauge invariant field strength kinetic term:

$$\mathcal{L} = \frac{\theta}{16\pi^2} F_{\mu\nu} \tilde{F}^{\mu\nu} \quad \text{where} \quad \tilde{F}_{\mu\nu} = \epsilon_{\mu\nu\rho\sigma} F^{\rho\sigma}. \tag{2.15}$$

Since this term can be written as a total derivative;

$$\mathcal{L} = \partial_\mu K^\mu \quad \text{where} \quad K_\mu = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma} A^\nu F^{\sigma\rho}, \tag{2.16}$$

the usual argument is that it only contributes a surface term to the action, so based on the boundary condition that the fields should go to zero (or more concretely, that the K_μ term goes to zero faster than the surface element diverges) in the limit that $r \to \infty$, this term can be neglected as it has no physical consequences. While this is certainly true in the abelian QED case, it does not hold in QCD. There, the total derivative has the form

$$K_\mu = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}(A^a_\nu G^a_{\sigma\rho} - \frac{2}{3}f_{abc}A^a_\nu A^b_\sigma A^c_\rho). \tag{2.17}$$

Hence, even if the field strength tensor $G^a_{\mu\nu}$ goes to zero rapidly enough, the non-Abelian A^3 term above means that K_μ might still not vanish at spatial infinity. In fact, there are field configurations that do not vanish. These $F\tilde{F}$ terms do not appear in a perturbative expansion of $\mathcal{L}_{\text{QCD}}$, however, they do have physical effects. To show this, we consider $\mathcal{L}_{\text{QCD}}$ with only two massless quarks:

$$\mathcal{L}_{\text{QCD}} = i\bar{Q}D_\mu\gamma^\mu Q + i\bar{u}_R D_\mu\gamma^\mu u_R + i\bar{d}_R D_\mu\gamma^\mu d_R. \tag{2.18}$$

Given that $m_{\{u,d\}} \ll \Lambda_{QCD}$, this is a fair approximation at low energy. At energies above Λ_{QCD}, the Lagrangian has a global 'chiral' $\text{SU}(2)_L \times \text{SU}(2)_R$ symmetry. Approaching Λ_{QCD}, this is spontaneously broken by QCD condensates $\langle \bar{q}q \rangle$ to a vectorial subgroup $\text{SU}(2)_V$. By Goldstone's theorem, there are three massless scalars associated with this symmetry breaking. These are identified with the pions. In fact the residual vectorial symmetry is explicitly broken by a small amount by electromagnetic interactions and the small $u - d$ mass splitting, which gives the pions a small mass ($\mathcal{O}(\Lambda_{\text{QCD}})$, though this is not calculable from first principles). The Lagrangian also has a $\text{U}(1)_A$ symmetry which is broken by QCD condensation. There is no Goldstone boson associated with this symmetry, however.

[**]See Ref. [164] for a pedagogical review.

Naïvely, this spontaneous breaking of the axial symmetry should result in the appearance of a pseudoscalar 0^- meson with mass of order the pion mass, but no such particle exists in the meson spectrum of QCD. The next such candidate is the η meson, but it is too heavy, as Weinberg showed that this particle can have a mass no greater than $\sqrt{3}m_\pi$ [165]. This was referred to as the $U(1)$ problem of QCD.

The resolution of this problem, due to 't Hooft [166, 167], was that non-perturbative gauge configurations known as *instantons* also contribute to the QCD vacuum, so that the full action is given by

$$S_{\text{QCD vacuum}} = \int d^4x \mathcal{L}_{\text{QCD}} + \frac{\theta g_s^2}{32\pi^2} \int d^4x \tilde{G}^a_{\mu\nu} G^{a,\mu\nu}. \tag{2.19}$$

That $m_\eta \gg m_\pi$ requires $\theta \neq 0$. However, other observables are sensitive to, and place strong bounds on, θ. It generates a contribution to the neutron electric dipole d_n for instance, $d_n \sim e\theta m_q/M_N^2$. Current bounds require $d_n \lesssim 3 \times 10^{-26} e$ cm [168]. This translates into the bound $\theta \lesssim 10^{-9}$. The question is then, why is the dimensionless parameter θ so small but apparently nonzero? There is no additional symmetry enhancement when θ is taken to zero, so this is not a technically natural small parameter. This is the strong-$\mathcal{CP}$ problem, and cannot be resolved within the Standard Model alone, therefore new physics, such as axions [169–172], is required to explain it.

2.3 The role of the top quark in specific BSM scenarios

2.3.1 Low energy supersymmetry

The most widely known solution to the hierarchy problem is supersymmetry. In general, supersymmetry is a postulate that the theory exhibits a symmetry under the transformation $Q|fermion\rangle = |boson\rangle$, $Q|boson\rangle = |fermion\rangle$, i.e. there is a unique transformation which maps each boson in the theory into a corresponding fermion and vice versa, thus keeping the overall theory invariant.

Clearly, the Standard Model does not exhibit this symmetry. This is linked to the problem of quadratic divergences in the Higgs mass. To see this, one can compute the explicit corrections to the Higgs mass due to loops of Standard Model particles. Focusing explicitly on the fermionic corrections (although the same argument applies for the W and Z bosons), due to N_f flavours of fermion, one finds [173]

$$m_H^2 = m_{H,tree}^2 + N_f^2 \frac{\lambda_f^2}{8\pi^2} \left(-\Lambda^2 + 6m_f^2 \log \frac{\Lambda}{m_f} - 2m_f^2 \right). \tag{2.20}$$

In a supersymmetric theory, for each fermion loop there would be a contribution due to a
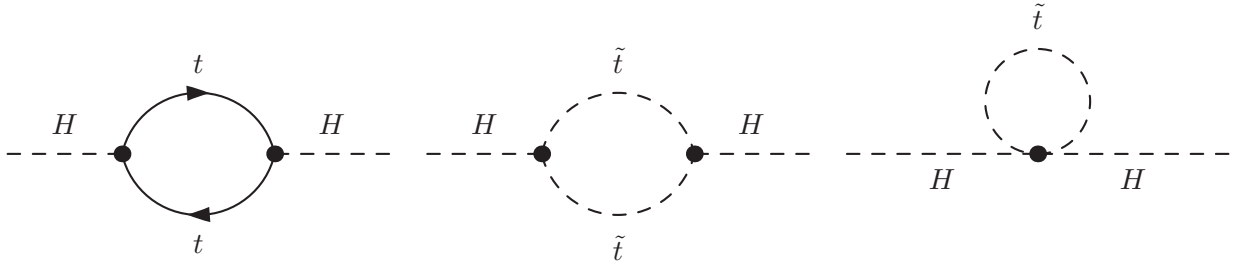
Figure 2.4: *One-loop diagrams contributing to the renormalisation of the Higgs mass in the Minimal Supersymmetric Standard Model: the correction due to a top quark loop (left) and corrections due to a loop of stop quarks $\tilde{t}$.*

corresponding scalar particle in the two diagrams on the right of Fig. 2.4, giving the mass correction

$$m_H^2 = m_{H,tree}^2 + N_s^2 \frac{\lambda_s}{16\pi^2} \left(-\Lambda^2 + 2m_s^2 \log \frac{\Lambda}{m_s} \right) - N_s^2 \frac{\lambda_s}{16\pi^2} v^2 \left(-1 + 2\log \frac{\Lambda}{m_s} \right). \quad (2.21)$$

Counting degrees of freedom, each fermion (a two-component Weyl spinor) must have two scalar 'partners', so $N_s = 2N_f$. If one makes the additional assumption that $\lambda_s = \lambda_f^2$, then upon adding these two contributions together, one finds the quadratic divergences in m_H^2 cancel entirely, leaving a logarithmic contribution to the Higgs mass

$$m_H^2 = m_{H,tree}^2 + N_f^2 \frac{\lambda_f^2}{4\pi^2} \left[(m_f^2 - m_s^2) \log \frac{\Lambda}{m_s} + 3m_f^2 \log \frac{m_s}{m_f} \right]. \quad (2.22)$$

If supersymmetry is exact, then $m_s = m_f$ and the corrections to the Higgs mass are exactly zero, to all orders in perturbation theory. This can be understood from symmetry grounds: a process involving virtual particles without their corresponding superpartners will violate supersymmetry and reintroduce divergences.

In practice, SUSY must be violated by some amount, otherwise the superpartners would have the same masses as their standard model counterparts, and would surely have been observed already. Therefore one must introduce SUSY breaking parameters to break the $m_s = m_f$ relation by some amount, though this amount cannot be too large otherwise the hierarchy problem will reappear. If SUSY is broken at the TeV scale, then it should show rich phenomenology at colliders such as the LHC. The top quark would play a central role here, both directly and indirectly.

To emphasise the latter, consider top pair production in the gg channel, with loop effects of scalar top (*stop*) quarks, as shown in Fig. 2.5. If the top pair production cross-section were very precisely measured, it could be used to place indirect bounds on the mass of the stop quark, and thus the scale of supersymmetry breaking. Indirect stop contributions
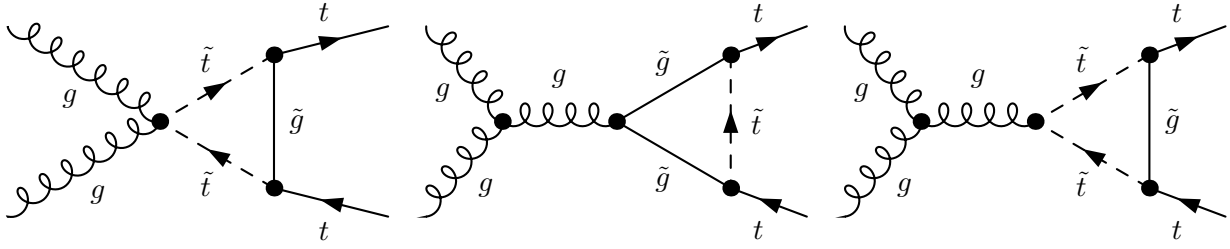
Figure 2.5: *Example one-loop diagrams contributing to top pair production in the gg channel, owing to the effects of virtual SUSY particles, namely the gluing g and the stop $\tilde{t}$.*

show up elsewhere, such as in the Higgs decay to $\gamma\gamma$, which is part mediated by a top loop [174]. This shows that the properties of the top quark are a valuable testing ground for the effects of new physics such as supersymmetry.

2.3.2 Little Higgs

The Higgs mass naturalness problem was solved in supersymmetry by adding extra degrees of freedom to cancel off the quadratic divergences, but more fundamentally this is due to extending the Poincaré algebra of spacetime symmetries. In general the Higgs mass can be made 'technically natural' by enlarging the SM symmetry group, such that quadratic corrections to μ^2 are forbidden by the extra symmetries. The 'Little Higgs' family [175] of models use this idea, by having the $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge group of the SM emerge from a spontaneously broken *global symmetry*. The gauge couplings of the SM break the residual symmetry explicitly by a small amount, and the Higgs emerges as one of the pseudo-Nambu-Goldstone bosons of this symmetry breaking. The pattern of the symmetry breaking ensures that the corrections to the Higgs mass are at most logarithmically sensitive to the cutoff Λ.

To illustrate this more concretely, we consider an extension of the minimal example of a global symmetry breaking $\mathrm{SU}(3)\to\mathrm{SU}(2)$ by the vev of a complex triplet field $\langle\Phi^T\rangle = \langle(\phi_1, \phi_2, \phi_3)\rangle = (0, 0, f)$. Instead, we consider two complex fields Φ_1 and Φ_2, each with its own set of 5 Nambu-Goldstone bosons, i.e. the symmetry breaking pattern is $[\mathrm{SU}(3)\to\mathrm{SU}(2)]^2$. To parameterise this symmetry breaking, Φ_1 and Φ_2 can be written as

$$\Phi_1 = e^{i\vec{\pi}_1/f_1}\begin{pmatrix}0\\0\\f_1\end{pmatrix}, \quad \Phi_2 = e^{i\vec{\pi}_2/f_2}\begin{pmatrix}0\\0\\f_2\end{pmatrix}. \tag{2.23}$$

For convenience we assume that the two vevs f_1 and f_2 are aligned. The Nambu-Goldstone bosons $\vec{\pi} = \pi^a T^a$ are given by the generators of $\mathrm{SU}(3)$ that are not also generators of $\mathrm{SU}(2)$.
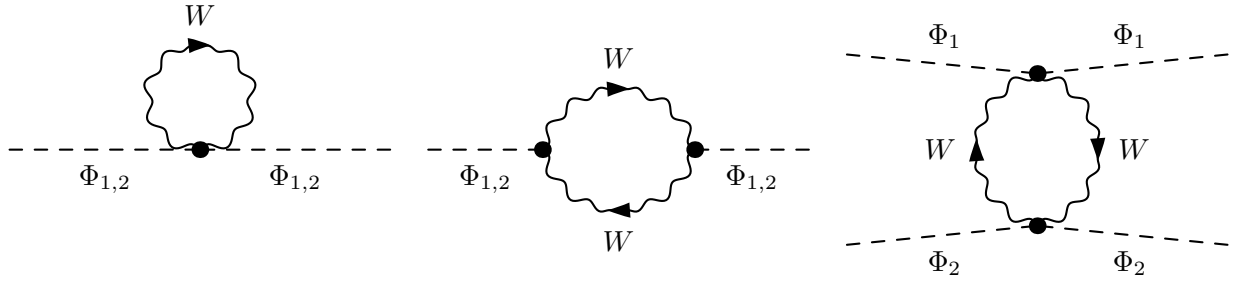
Figure 2.6: *Quadratically (left) and logarithmically (centre) divergent one-loop diagrams from Eq. (2.25) that do not renormalise the Higgs mass, and logarithmically divergent 1-loop diagrams that do renormalise the Higgs mass (right).*

We can write explicitly, for each set $\vec{\pi}$

$$\pi^a T^a = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & \\ 0 & 0 & H \\ H & & 0 \end{pmatrix} + \frac{\eta}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{2.24}$$

Four of the Goldstones are in the complex doublet H, the remaining one is in the singlet η, which can be ignored for our purposes. The factors in front of the Goldstone matrices ensure that the H and η kinetic terms are canonically normalised. As well as the Goldstones, there are also massive radial excitations r, which are assumed to be heavy so that they are integrated out. It is also assumed for simplicity that there are degenerate symmetry breaking scales $f_1 = f_2 = f$. The Lagrangian for this toy model is then

$$\mathcal{L}_{\mathrm{LH}} = |D_\mu \Phi_1^2|^2 + |D_\mu \Phi_2^2|^2. \tag{2.25}$$

Upon expanding out the Lagrangian, one generates the 1-loop correlation functions corresponding to the two diagrams on the left hand side of Fig. 2.6:

$$\mathcal{A}_{1-\mathrm{loop}} \sim \frac{g^2 \Lambda^2}{16\pi^2}(\Phi_1^\dagger \Phi_1 + \Phi_2^\dagger \Phi_2) = \frac{g^2 \Lambda^2}{16\pi^2}(2f^2). \tag{2.26}$$

In addition, one generates the term corresponding to the diagram on the right of Fig. 2.6.

$$\mathcal{A}_{1-\mathrm{loop}} \sim \frac{g^4}{16\pi^2} \log\left(\frac{\Lambda^2}{\mu^2}\right) |\Phi_1^\dagger \Phi_2|^2 = \frac{g^4}{16\pi^2} \log\left(\frac{\Lambda^2}{\mu^2}\right)\left(f^2 - 2H^\dagger H + \frac{(H^\dagger H)^2}{f^2} + \ldots\right). \tag{2.27}$$

Hence, the 1-loop corrections to the Higgs mass term are at most logarithmically divergent. To see why this has happened, let us focus on the gauge part of $\mathcal{L}_{\mathrm{LH}}$:

$$\mathcal{L}_{\mathrm{LH}} = |g_1^2 A_\mu \Phi_1|^2 + |g_2^2 A_\mu \Phi_2|^2. \tag{2.28}$$

This explicitly breaks the SU(3)×SU(3) global symmetry to a gauged diagonal subgroup: so only one of the SU(3)→SU(2) breaking mechanisms is exact. The other is explicitly broken by a small amount, giving its Goldstone bosons a small mass, one of which we then take to be the Higgs. Setting either g_1 or g_2 to zero restores the full $[SU(3)]^2$ symmetry. In the case of $g_2 = 0$, for example, we have two independent symmetries

$$\Phi_1 \to U_1 \Phi_1, \quad A_\mu \to U_1 A_\mu U_1^\dagger, \quad \Phi_2 \to U_2 \Phi_2, \tag{2.29}$$

whereas in the case of $g_1 = 0$ we have the symmetries

$$\Phi_1 \to U_1 \Phi_1, \quad A_\mu \to U_2 A_\mu U_2^\dagger, \quad \Phi_2 \to U_2 \Phi_2. \tag{2.30}$$

So when either of the gauge couplings is set to zero, π is an exact Nambu-Goldstone boson, so the corrections to its mass can only be proportional to $g_1 g_2$, i.e. the exact symmetry can only be *collectively* broken by the two fields. There are no quadratically divergent diagrams involving g_1 and g_2 at 1-loop, however, so the Higgs mass is stabilised at this order by this *collective symmetry breaking* [176, 177].

The same trick for the gauge loops can be played for the top quark loops. To ensure that there is collective symmetry, one enlarges the quark doublets into triplets; $Q_L \to \Psi_L = (t_L, b_L, T_L)$, by adding an extra fermionic partner for each generation. One finds again that the quadratic divergences due to the top quark loop are cancelled by the top partner T. It is this extra top partner that has an impact on top quark phenomenology. It will in general mix with the top quark, so it can be produced via Wb fusion, Depending on its quantum numbers, it may decay via $T \to th$, $T \to tZ$, $T \to bW$ [178], which would lead to large enhancements of cross sections for top quarks associated with electroweak and Higgs bosons.

2.3.3 Warped extra dimensions

An alternative approach to explaining the large hierarchy between the electroweak scale and the Planck mass is through warped extra dimensions. The most studied scenario is the Randall-Sundrum (RS) model [179, 180], where the Standard Model field content resides on the 4-dimensional boundary of a 5-dimensional bulk, in which the gravitational degrees of freedom propagate. The extra dimension is compactified onto a $S^1/\mathbb{Z}_2$ orbifold (a circle with an additional $\mathbb{Z}_2$ symmetry $\phi = -\phi$) of radius r_c, and the two boundaries of the 5D bulk are taken to be at the points $\phi = 0$ and $\phi = \pi$, where a 4-dimensional field theory resides, such that there is a *visible* boundary or brane and a *hidden* one. The full 5D metric

$G_{MN}(x^\mu, \phi)$ is then related to the 4D metrics by the boundary conditions:

$$g^{vis}_{\mu\nu} = G_{\mu\nu}(x^\mu, \phi = \pi), \quad g^{hid}_{\mu\nu} = G_{\mu\nu}(x^\mu, \phi = 0). \tag{2.31}$$

The Einstein-Hilbert action for the 5D theory is then given by

$$S = S_{grav} + S_{vis} + S_{hid} = \int d^4x \int_{-\pi}^{\pi} d\phi \sqrt{-G}(\Lambda + 2M^3R) + \int d^4x (\sqrt{-g_{vis}}\mathcal{L}_{\text{vis}} + \sqrt{-g_{hid}}\mathcal{L}_{\text{hid}}), \tag{2.32}$$

where Λ is a cosmological constant, M is a universal mass scale extracted to give the field ϕ the same units as in the 4D theory and R is the Ricci scalar. A solution to the Einstein field equations for the above action is

$$ds^2 = e^{-2kr_c|\phi|}\eta_{\mu\nu}dx^\mu dx^\nu + r_c^2 d\phi^2. \tag{2.33}$$

This non-factorisable metric describes flat 4D spacetime modified by an exponential *warp factor* $2kr_c|\phi|$. The parameter k is a scale relating the 'observed' 4D Planck scale to the Planck scale in the bulk. The radius r_c describes the compactification of the 5th dimension, and may be taken to be near the Planck length, if one views the model as originating from a string/M-theory UV completion. To see how this is relevant for low energy physics, we can expand the metric about its local fluctuations $g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}(x)$, substitute in a Higgs field into $\mathcal{L}_{\text{vis}}$ in the action, and perform a wave-function renormalisation so that the Higgs kinetic term is canonically normalised, and one finds that any mass or vev m_0 in the fundamental 5D theory is related to the mass/vev in the visible theory by

$$m_{vis} = e^{-kr_c\pi}m_0. \tag{2.34}$$

So a hierarchy of order $v/M_{Pl} \sim 10^{-16}$ in the visible theory translates into a hierarchy of size $kr_c \sim 12$ in the fundamental theory, thus it is much more natural for the bulk curvature k to live near the Planck scale $1/r_c$. This is a compelling solution of the hierarchy problem, because it shows that weak scale masses can be determined by parameters not far from the Planck scale, but in a natural way. Why is this relevant for top quark physics?

The main phenomenological prediction of warped extra dimension models is that, because the extra dimension has periodic (Dirichlet) boundary conditions, the fields that are allowed to propagate in the bulk will have an infinite tower of *Kaluza-Klein* modes of mass m_n, analogous to standing wave modes on a closed string[††], which can be coupled to the SM fields that also propagate in the bulk [185, 186], and which, according to Eq. (2.34) can have masses near the electroweak scale. Moreover, the KK modes will couple most

---

[††]Kaluza-Klein excitations are also present in other theories of extra dimensions, such as the ADD models [181, 182] and universal extra dimensions [183, 184].

strongly to fields that are localised near the $\phi = \pi$ IR brane, and most weakly coupled to fields localised on the UV brane.

To keep a hierarchy between v and M_{Pl}^{4D}, the Higgs field must be localised on the IR brane ($\phi = \pi$), cf Eq. (2.34). However, the large top Yukawa means there must be a strong overlap between the top and Higgs wave-functions, so the top most also be located close to the IR [187, 188]. Hence, the top is expected to couple strongly to the new Kaluza-Klein modes of the 5th dimension, which motivates the search for heavy resonances of spin-1 (the KK modes of the γ, g, W and Z) and spin-2 (the KK modes of the massless graviton), decaying into $t\bar{t}$ pairs.

The preceding section provides a strong argument for processes involving top quarks as a well-motivated place to look for BSM physics at colliders, albeit in the limited context of specific models. The remainder of this chapter will turn to a more model-independent formulation of the effects of new particles and couplings, which views the Standard Model as the first part of an effective field theory, and discusses the sector of this effective theory that may be probed at hadron colliders.

2.4 Principles of effective field theory

2.4.1 The Euler-Heisenberg Lagrangian

As a warmup, we consider an example from electromagnetism. Imagining a different Universe where the electron was much heavier and had not yet been detected directly, so that the 'full' Lagrangian for electromagnetism is just that of the free Maxwell theory.

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{2.35}$$

If the electromagnetic process of four-photon scattering had been experimentally observed, it could not be described by this Lagrangian. The lowest-order Lagrangian that could describe this process would have to have four field-strength tensors, since the Maxwell theory is abelian. It should also respect Lorentz symmetry and U(1) gauge symmetry, and can be written as

$$\mathcal{L}_{4\gamma} = A(F_{\mu\nu}F^{\mu\nu})^2 + B(F_{\mu\nu}\tilde{F}^{\mu\nu})^2. \tag{2.36}$$

Since the Lagrangian must have mass dimension 4 (we work in natural units), the coefficients A and B must have mass dimension -4. We can then write

$$\mathcal{L}_{4\gamma} \equiv \frac{1}{\Lambda^4}\left(c_1(F_{\mu\nu}F^{\mu\nu})^2 + c_2(F_{\mu\nu}\tilde{F}^{\mu\nu})^2\right). \tag{2.37}$$

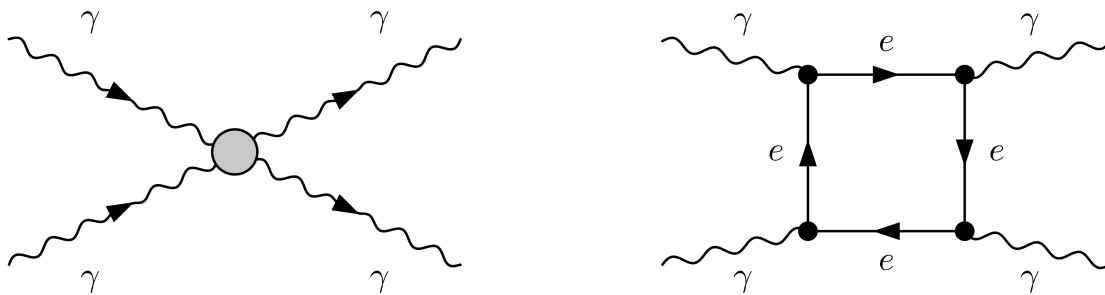The coefficients c_1 and c_2 are dimensionless, and Λ is some generic mass scale. From

Figure 2.7: $\gamma\gamma \to \gamma\gamma$ *scattering at tree-level in the Euler-Heisenberg effective theory (left) and at one-loop level in quantum electrodynamics (right).*

this Lagrangian we know that the scattering cross-section σ for $\gamma\gamma \to \gamma\gamma$ scales as $1/\Lambda^8$. However, since σ is just an area, it has units of $[mass]^{-2}$, so there must be some other mass scale in the problem. The only other mass scale in the problem, however, is the frequency of the incoming photons $\omega^{\ddagger\ddagger}$, so we can say that the cross-section scales as

$$\sigma \sim \frac{\omega^6}{\Lambda^8} + \dots \tag{2.38}$$

The power of effective field theory is manifest: by general arguments of symmetry and making no assumptions about the nature of the underlying interaction, one can make powerful deductions about the scaling behaviour of scattering processes in quantum field theory.

There is a caveat: the ellipsis denotes higher-order corrections due to neglected operators of dimension $D > 8$, the leading term of which will be $\mathcal{O}(\omega^8/\Lambda^{10})$. There are an infinite number of such operators, so these corrections must be small in order for their omission to be valid. In other words, the validity of the effective theory requires that $\omega \ll \Lambda$. But in the pure effective theory, Λ is a free parameter, so there is no *a priori* guarantee that this is true. In reality, of course there is a full theory of electromagnetism: quantum electrodynamics, so one can compute the scattering amplitude in terms of its parameters: the fine-structure constant α and the electron mass m_e, corresponding to the Feynman diagram on the right of Fig. 2.7, and match them onto the EFT parameters. One finds [189]

$$\Lambda = \frac{m_e}{\sqrt{\alpha}}, \quad c_1 = \frac{1}{90}, \quad c_2 = \frac{7}{90}. \tag{2.39}$$

Therefore, provided the photon scattering frequencies $\omega \ll m_e/\sqrt{\alpha} \sim \mathcal{O}(\mathrm{MeV})$, this treatment is valid, and the loop diagram of Fig. 2.7 does not have to be calculated.

---

‡‡For simplicity we assume them to be of the same order of magnitude.

2.4.2 Fermi theory of weak decay

The textbook example of an effective field theory is the Fermi low-energy theory of weak decay. One can arrive at the theory by starting from the full electroweak model and integrating out the heavy degrees of freedom, i.e. the W and Z boson. In the electroweak theory, nuclear β decay is mediated by the transition $d \to uW \to ul\nu$. This has the transition amplitude [190]

$$\mathcal{A}_{\text{full}} = \left(\frac{ig}{\sqrt{2}}\right)^2 (\bar{u}\gamma_\mu P_L d)(\bar{l}\gamma_\nu P_L \nu)\left(\frac{-ig^{\mu\nu}}{p^2 - M_W^2}\right), \qquad (2.40)$$

where P_L is the left-handed projection operator $(1 - \gamma_5)/2$. The low energy continuum is obtained by Taylor expanding the W boson propagator in the limit $p^2 \ll M_W^2$.

$$\left(\frac{-ig^{\mu\nu}}{p^2 - M_W^2}\right) = \frac{-ig^{\mu\nu}}{M_W^2}\left(1 + \frac{p^2}{M_W^2} - \frac{p^4}{M_W^4} + \dots\right). \qquad (2.41)$$

Hence the higher order terms decouple rapidly in this limit, so that the amplitude in the low energy theory is given by the first term in the series, multiplied by a $D = 6$ four-fermion contact term:

$$\mathcal{A}_{\text{EFT}} = \left(\frac{g^2}{2M_W^2}\right)(\bar{u}\gamma_\mu P_L d)(\bar{l}\gamma^\mu P_L \nu). \qquad (2.42)$$

In fact this is the same β decay amplitude that was written down by Fermi in his contact interaction model, though he wrote in terms of the nucleon wavefunctions $u \to p$ and $d \to n$, and omitted the projection operator P_L as parity violation had not yet been observed. He also parametrised it in terms of an overall dimensionful coupling G_F, which we can obtain an expression for by matching the electroweak parameters to the parameters of the Fermi theory, giving

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2}, \qquad (2.43)$$

with the value of G_F given in the first chapter. Provided the condition $p^2 \ll M_W^2$ is justified, one can calculate to a good approximation all weak scattering processes, such as muon decay and meson mixing, without knowing the details of the gauge structure of the underlying electroweak theory. In fact, if one makes assumptions about the underlying couplings, one can even predict the regime of validity of the EFT. Setting $g = 1$ and using the measured value of G_F, and neglecting $\mathcal{O}(1)$ coefficients above, for example, one finds $M_W \approx 290$ GeV, not far from its actual value of 80 GeV, and very close to the electroweak scale $v \sim 246$ GeV where the EFT would no longer be valid. This highlights another strength of EFT: using the measurements of the low energy parameters, one can infer details about some of the high-energy ones, by making broad assumptions about the perturbativity of the underlying theory. This *matching* procedure will be returned to later

in the thesis.

2.4.3 Renormalisation group treatment

Here we perform a more systematic analysis of how the dimensionality of an operator determines at which scale it becomes relevant, by means of the renormalisation group. To isolate the behaviour of operators of a certain dimensionality when the renormalisation scale is varied, let us return to our prototypical example of a scalar field in four dimensions with a quartic self-interaction. This time, we supplement the Lagrangian with an infinite series of higher-dimensional (i.e. $D > 4$) operators [191].

$$\mathcal{L}_{\mathrm{E}} = \frac{1}{2}(\partial\phi)^2 + \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4!}\phi^4 + \sum_n \left(\frac{c_n}{\Lambda^{2n}}\phi^{4+2n} + \frac{d_n}{\Lambda^{2n}}(\partial\phi)^2\phi^{2+2n} + \dots \right), \qquad (2.44)$$

where the Lagrangian has been rotated to imaginary time in order to perform the path integral, defining a 'Euclidean action'. The ellipsis denotes all operators with higher derivatives. Demanding that the effective theory preserve the original $\phi \to -\phi$ symmetry eliminates terms of odd mass dimension. The scale Λ has been introduced to keep the coupling constants c_i and d_i dimensionless. The kinetic and mass terms both have dimension $D = 2$, the quartic coupling has $D = 4$. The question is then: how does the dimensionality of an operator influence its renormalisation scaling behaviour?

To isolate field configurations that are most relevant at a certain momentum/length, we can perform the path integral

$$\int D\phi e^{-S_E} \quad \text{where} \quad S_E = \int d^4x \mathcal{L}_{\mathrm{E}}. \qquad (2.45)$$

We can consider a field configuration $\tilde{\phi}$ in the path integral with amplitude ϕ_k, and wavenumber k_μ, i.e. a 'wavelet' that is confined to a spacetime volume $L^4 = (2\pi/k)^4$. In momentum space, the Euclidean action can be trivially obtained by Fourier transforming Eq. (2.44). Then the action is given by

$$S_E = (2\pi)^4 \left[\frac{\hat{\phi}_k^2}{2} + \frac{m^2}{k^2}\hat{\phi}_k^2 + \frac{\lambda}{4!}\hat{\phi}_k^4 + \sum_n \left(c_n \left(\frac{k^2}{\Lambda^{2n}}\right)^2 \hat{\phi}_k^{4+2n} + d_n \left(\frac{k^2}{\Lambda^{2n}}\right)^2 \hat{\phi}_k^{4+2n} + \dots \right) \right],$$
$$(2.46)$$

where $\hat{\phi}_k \equiv \phi_k/k$. The contribution of this single mode to the path integral is

$$\int d\hat{\phi}_k e^{-S_E}. \qquad (2.47)$$

Clearly it will be dominated by values of $\hat{\phi}_k$ for which $S_E \lesssim 1$. As the amplitude $\hat{\phi}_k$ becomes

large, then the kinetic term $(2\pi)^4 \hat{\phi}_k^2/2$ will dominate the action, i.e the path integral will get its dominant contribution for $\phi_k \sim k/(2\pi^2)$.

When k is decreased, the higher-dimensional terms proportional to c_i and d_i get smaller and smaller; they are called *irrelevant* operators; their effects become increasingly decoupled as we move from the ultraviolet to the infrared. The mass term, on the other hand, becomes increasingly dominant; it is a *relevant* operator. The quartic operator is neither relevant nor irrelevant, its effects are of apparently equal strength for small and large k; it is referred to as a *marginal* operator.

Another way of deriving the scaling properties of these operators (the method originally employed by Wilson [192–194]) is to consider a random field configuration $\phi(x)$, and look at how its corresponding action changes when we perform a passive transformation $\phi(x) \to \phi(\xi x)$, i.e. when we move across different length scales. If we just consider a plane wave, for instance, then the transformation is given by $\phi(\xi x) = e^{i\xi k \cdot x}$, so that the limit $\xi \to \infty$ corresponds to shorter wavelengths $k' = \xi k$. Then the action becomes

$$
\begin{aligned}
S_E(\phi(\xi x); \Lambda, m^2, \lambda, c_n, d_n) &= \int d^4 x \frac{1}{2}(\partial\phi(\xi x))^2 + \frac{1}{2}m^2\phi(\xi x)^2 + \frac{\lambda}{4!}\phi(\xi x)^4 \\
&\quad + \sum_n c_n \frac{\phi^{4+2n}(\xi x)}{\Lambda^{2n}} + \sum_n d_n \frac{(\partial\phi(\xi x))^2 \phi^{2n}(\xi x)}{\Lambda^{2n}} \\
&= \int d^4 x' \frac{1}{2}(\partial'\phi'(x'))^2 + \frac{1}{2}m^2\xi^{-2}\phi'(x')^2 + \frac{\lambda}{4!}\phi'(x')^4 \\
&\quad + \sum_n \left(c_n \xi^{2n} \frac{\phi'^{4+2n}(x')}{\Lambda^{2n}} + d_n \xi^{2n} \frac{(\partial'\phi'(x'))^2 \phi'^{2n}(x')}{\Lambda^{2n}} \right),
\end{aligned}
\tag{2.48}
$$

where $\phi'(x) = \xi^{-1}\phi(x)$, and $x' = \xi x$. But since we integrate over x and x', we can just compare the integrands directly. Relabelling the dummy variable $x' \to x$ shows that our transformation returns the original action, but with rescaled fields and couplings.

$$
S_E(\phi(\xi x); \Lambda, m^2, \lambda, c_n, d_n) = S_E(\xi^{-1}\phi(x); \xi^{-2}m^2, \lambda, c_n\xi^{2n}, d_n\xi^{2n}),
\tag{2.49}
$$

so the rescaled fields and couplings are:

$$
\phi \to \xi^{-1}\phi, \quad m^2 \to \xi^{-2}m^2, \quad \lambda \to \lambda, \quad c_n \to \xi^{2n}c_n, \quad d_n \to \xi^{2n}d_n.
\tag{2.50}
$$

In the infrared limit $\xi \to 0$, it can be seen that the mass term becomes increasingly important, the higher-dimensional operators c_n and d_n become increasingly irrelevant, and the quartic and kinetic terms stay marginal. This result follows purely from dimensional analysis, and did not rely on any unique symmetry properties of scalar fields. We can then

make the general statement that there are three types of operator scaling behaviour in a four-dimensional quantum field theory.

- **Relevant operators:** Operators $\mathcal{O}_i^{(d_i)}$ with dimension $d_i < 4$. Their effects become increasingly large at low energies, and increasingly decoupled at high energies.

- **Marginal operators:** Operators $\mathcal{O}_i^{(d_i)}$ with dimension $d_i = 4$. Their effects are naïvely the same across all scales, i.e. they appear to be conformal. However they may scale logarithmically, and a full radiative calculation is needed to obtain their scaling behaviour.

- **Irrelevant operators:** Operators $\mathcal{O}_i^{(d_i)}$ with dimension $d_i > 4$. Their effects become increasingly large at high energies, and increasingly decoupled in the infrared.

It was once considered a miracle that the Standard Model contained only marginal and relevant (and thus renormalisable) operators. From a modern perspective, we know that there is no miracle and this follows completely from the renormalisation group: the Standard Model is a theory of low energy physics (compared to the Planck scale), therefore the higher-dimensional operators would not be expected, because they only 'switch on' as we move towards the cutoff Λ for the Standard Model. If there exist heavy new degrees of freedom, then at some scale between the electroweak and Planck scales their effects on electroweak scale observables can be described generally by supplementing the Standard Model Lagrangian with operators of dimension > 4. These operators are built completely out of Standard Model fields since the underlying new heavy degrees of freedom have been integrated out. This is the Standard Model effective field theory (SMEFT), and it thus provides a completely model-independent way of searching for the effects of unknown heavy new physics on electroweak scale observables. In the next section we will discuss the SMEFT in detail.

2.5 The Standard Model effective field theory

To derive the effective Lagrangian for the Standard Model [195–197], all we have to do is write down the expansion in powers of c_n/Λ^n, as in Eq. (2.44), but with the full SM field content, rather than just one scalar with a quartic interaction. This time we do not have a $\mathbb{Z}_2$ symmetry to respect, so odd powers of Λ^{-1} are allowed. The effective Lagrangian can then be written

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{c_i^{(5)} \mathcal{O}_i^{(5)}}{\Lambda} + \sum_i \frac{c_i^{(6)} \mathcal{O}_i^{(6)}}{\Lambda^2} + \dots, \tag{2.51}$$

where the sum is over the full operator set at each mass dimension and the ellipsis denotes all operators at $D \geq 7$. The full set of operators can be derived systematically, simply by

using dimensional analysis to write down the list of operators of a given dimension that respect the symmetry constraints, i.e. Lorentz invariance and the full $SU(3)_C \times SU(2)_L \times U(1)_Y$ SM gauge symmetry. Then, care must be taken to ensure there are no redundancies in the operator set; that is, each operator generates a unique contribution to the S-matrix that cannot be expressed in terms of other operators. We will derive the dimension-five SM Lagrangian as an illustrative example.

2.5.1 $D = 5$

To recap, the SM Lagrangian is composed entirely of spin-$\frac{1}{2}$ fermion fields of dimension 3/2, scalar and vector fields of dimension 1, and field strength tensors of dimension 2, as well as various covariant derivative operators. At dimension 5, then, naïvely there are several types of operator that can be constructed, but more careful analysis shows that most are forbidden for symmetry reasons:

- Clearly, no fermion-only operators are allowed, because 5 is not a multiple of $\frac{3}{2}$. Scalar only operators are also forbidden, because the Higgs only appears in doublets, so there must be an even number of scalars in the term.

- An operator with two fermions and two scalars is allowed dimensionally. Two combinations of scalars would be allowed, either $(\varphi^\dagger \varphi)$ or $(\varphi \varphi)$. The first case requires that the two fermions must also combine to hypercharge zero, so they must be Hermitian conjugates of the same fermion multiplet: $\bar{\psi}\psi$, which vanishes for chiral fermions. The second case is allowed, provided the scalars multiply to give an SU(2) triplet (the singlet product of two equal doublets is zero). Then the two fermions must also form a triplet to dot this into a scalar. Each fermion must then be an SU(2) doublet. The term can then be written as

$$\mathcal{L}_5 = \epsilon_{ij} \bar{L}_i \varphi_j \epsilon_{kl} L_k \varphi_l + h.c. \tag{2.52}$$

where $i, j \in \{1, 2\}$ etc. denote weak isospin indices.

In fact, this is the only allowed dimension-five operator [198] in the Standard Model effective theory. No analogous operator may be formed with quark fields, since replacing $L \to Q$ does not give a colour singlet. Operators with two vector fields and two fermions are forbidden because the fermion bilinear must be a hypercharge zero SU(2) singlet $\bar{\psi}\psi$, and so vanishes for chiral fermions. Other combinations involving vector fields cannot be constructed on dimensional grounds. Expanding Eq. (2.52) after electroweak symmetry breaking generates a Majorana-like mass term for the neutrino $m_\nu \bar{\nu}_L \nu_L^C$, and mixing between the neutrino flavour eigenstates. It is therefore required to be non-zero by neutrino phenomenology.

However, the mass terms are proportional to $m_\nu \sim v^2/\Lambda$, which points to $\Lambda \sim 10^{13}$ GeV based on current neutrino mass limits. It is therefore not accessible at collider energies, and not explored in the remainder of this thesis.

2.5.2 $D = 6$

At dimension-six, many more operators are allowed[§§]. Deriving the full, non-redundant dimension-six operator set is somewhat more involved, so it will not be fully reproduced here, we will merely comment on some of its features. All the operators are built out of the same objects: field strength tensors of dimension two (shorthanded as X), Higgs doublets of dimension one (denoted φ), fermion fields ψ of dimension $\frac{3}{2}$ and various covariant derivatives D of dimension one. By simple power-counting, we can denote the operators as belonging to one of three classes:

Bosonic operators: These contain no fermion fields. There must be an even number of Higgs doublets, and an even number of covariant derivative operators, to ensure that all Lorentz indices are contracted. The allowed combinations are then X^3, $X^2\varphi^2$, X^2D^2, $X\varphi^4$, XD^4, $X\varphi^2D^2$, φ^6, φ^4D^2 and φ^4D^2. We can eliminate several of these classes. Firstly, Lorentz symmetry forbids $X\varphi^4$ terms, which are not Lorentz contracted. All XD^4 terms can also be moved to X^2D^2 terms by use of the identity $[D_\mu, D_\nu] \sim X_{\mu\nu}$. As for the remaining classes:

$\underline{\varphi^2D^4}$: By the equations of motion:

$$
\begin{aligned}
(D_\mu D^\mu \varphi)^j &= m^2\varphi^j - \lambda(\varphi^\dagger\varphi)\varphi^j - \bar{e}y_e^\dagger L^j + \epsilon_{jk}\bar{Q}^k y_u u - \bar{d}y_d^\dagger Q^j, \\
(D^\rho G_{\rho\mu})^A &= g_s(\bar{Q}\gamma_\mu T^A Q + \bar{u}\gamma_\mu T^A u + \bar{d}\gamma_\mu T^A d), \\
(D^\rho W_{\rho\mu})^I &= \frac{g}{2}\left(\varphi^\dagger i \overset{\leftrightarrow}{D}{}^I_\mu \varphi + \bar{L}\gamma_\mu \tau^I L + \bar{Q}\gamma_\mu \tau^I Q\right), \\
\partial^\rho B_{\rho\mu} &= g'Y_\varphi \varphi^\dagger i \overset{\leftrightarrow}{D}_\mu \varphi + g'\sum_i Y_\psi \bar{\psi}\gamma_\mu \psi.
\end{aligned}
\tag{2.53}
$$

these can be moved to operators in the bosonic classes φ^2XD^2 and φ^4D^2, as well as the single-fermionic current operator class $\psi^2\phi D^2$.

$\underline{\varphi^2XD^2}$: The identity $[D_\mu, D_\nu] \sim X_{\mu\nu}$ again moves some operators to the φ^2X^2 class. Also, using the equations of motion for the gauge field and the Bianchi identity $D_{[\rho}X_{\mu\nu]} = 0$, the remaining operators of this class are moved either to the bosonic class φ^4D^2 or the fermionic class $\psi^2\phi D^2$.

$\underline{X^2D^2}$: All operators of this class can be reduced to operators in the class X^3, φ^2XD^2, or ψ^2XD, or made to vanish by the equations of motion.

[§§]The elimination of all the redundancies of the original operator set written down in Ref. [197] to the basis of Ref. [199] was partially done in several intermediate papers [200–207].

X^3		φ^6 and $\varphi^4 D^2$		$\psi^2\varphi^3$	
$\mathcal{O}_G$	$f_{ABC}G_\mu^{A,\nu}G_\nu^{B,\rho}G_\rho^{C,\mu}$	$\mathcal{O}_\varphi$	$(\varphi^\dagger\varphi)^3$	$\mathcal{O}_{e\varphi}$	$(\varphi^\dagger\varphi)(\bar{L}e\varphi)$
$\mathcal{O}_{\tilde{G}}$	$f_{ABC}\tilde{G}_\mu^{A,\nu}G_\nu^{B,\rho}G_\rho^{C,\mu}$	$\mathcal{O}_{\varphi\Box}$	$(\varphi^\dagger\varphi)\Box(\varphi^\dagger\varphi)$	$\mathcal{O}_{u\varphi}$	$(\varphi^\dagger\varphi)(\bar{Q}u\tilde{\varphi})$
$\mathcal{O}_W$	$\epsilon^{IJK}W_\mu^{I,\nu}W_\nu^{J,\rho}W_\rho^{K,\mu}$	$\mathcal{O}_{\varphi D}$	$(\varphi^\dagger D^\mu\varphi)^\star(\varphi^\dagger D_\mu\varphi)$	$\mathcal{O}_{d\varphi}$	$(\varphi^\dagger\varphi)(\bar{Q}d\varphi)$
$\mathcal{O}_{\tilde{W}}$	$\epsilon^{IJK}\tilde{W}_\mu^{I,\nu}W_\nu^{J,\rho}W_\rho^{K,\mu}$				

$X^2\varphi^2$		$\psi^2 X\varphi$		$\psi^2\varphi^2 X$	
$\mathcal{O}_{\varphi G}$	$\varphi^\dagger\varphi G_{\mu\nu}^A G^{A,\mu\nu}$	$\mathcal{O}_{eW}$	$(\bar{L}\sigma^{\mu\nu}e)\tau^I\varphi W_{\mu\nu}^I$	$\mathcal{O}_{\varphi l}^{(1)}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar{L}\gamma^\mu L)$
$\mathcal{O}_{\varphi\tilde{G}}$	$\varphi^\dagger\varphi\tilde{G}_{\mu\nu}^A G^{A,\mu\nu}$	$\mathcal{O}_{eB}$	$(\bar{L}\sigma^{\mu\nu}e)\varphi B_{\mu\nu}$	$\mathcal{O}_{\varphi l}^{(3)}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu^I\varphi)(\bar{L}\tau^I\gamma^\mu L)$
$\mathcal{O}_{\varphi W}$	$\varphi^\dagger\varphi W_{\mu\nu}^I W^{I,\mu\nu}$	$\mathcal{O}_{uG}$	$(\bar{Q}\sigma^{\mu\nu}T^A u)\tilde{\varphi}G_{\mu\nu}^A$	$\mathcal{O}_{\varphi e}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar{e}\gamma^\mu e)$
$\mathcal{O}_{\varphi\tilde{W}}$	$\varphi^\dagger\varphi\tilde{W}_{\mu\nu}^I W^{I,\mu\nu}$	$\mathcal{O}_{uW}$	$(\bar{Q}\sigma^{\mu\nu}u)\tau^I\tilde{\varphi}W_{\mu\nu}^I$	$\mathcal{O}_{\varphi Q}^{(1)}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar{Q}\gamma^\mu Q)$
$\mathcal{O}_{\varphi B}$	$\varphi^\dagger\varphi B_{\mu\nu} B^{\mu\nu}$	$\mathcal{O}_{uB}$	$(\bar{Q}\sigma^{\mu\nu}u)\tilde{\varphi}B_{\mu\nu}$	$\mathcal{O}_{\varphi Q}^{(3)}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu^I\varphi)(\bar{Q}\tau^I\gamma^\mu Q)$
$\mathcal{O}_{\varphi\tilde{B}}$	$\varphi^\dagger\varphi\tilde{B}_{\mu\nu} B^{\mu\nu}$	$\mathcal{O}_{dG}$	$(\bar{Q}\sigma^{\mu\nu}T^A d)\varphi G_{\mu\nu}^A$	$\mathcal{O}_{\varphi u}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar{u}\gamma^\mu u)$
$\mathcal{O}_{\varphi WB}$	$\varphi^\dagger\tau^I\varphi W_{\mu\nu}^I B^{\mu\nu}$	$\mathcal{O}_{dW}$	$(\bar{Q}\sigma^{\mu\nu}d)\tau^I\varphi W_{\mu\nu}^I$	$\mathcal{O}_{\varphi d}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar{d}\gamma^\mu d)$
$\mathcal{O}_{\varphi\tilde{W}B}$	$\varphi^\dagger\tau^I\varphi\tilde{W}_{\mu\nu}^I B^{\mu\nu}$	$\mathcal{O}_{dB}$	$(\bar{Q}\sigma^{\mu\nu}d)\varphi B_{\mu\nu}$	$\mathcal{O}_{\varphi ud}$	$(\tilde{\varphi}^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar{u}\gamma^\mu d)$

Table 2.1: *The non-redundant bosonic and single fermionic-current $D = 6$ operators in the 'Warsaw basis' described here. For readability we do not explicitly display the fermion generation indices, but where relevant they are denoted by an extra superscript. For example $\mathcal{O}_{e\varphi}^{23} = (\varphi^\dagger\varphi)(\bar{\mu}, \bar{\nu}_\mu)\tau\varphi$.*

The only surviving bosonic operator classes are then $X^3, X^2\varphi^2, \varphi^6$ and $\varphi^4 D^2$. The non-redundant operators in each of these classes are shown in columns 1, 2 and 4 of Tab. 2.1.

Single-fermionic current operators: The classes allowed here are: $\psi^2 D^3$, $\psi^2\varphi D^2$, $\psi^2 XD$, $\psi^2\phi^3$, $\psi^2 X\varphi$ and $\psi^2\phi^2 D$. We can make use of the following equations of motion for the fermion currents.

$$i\slashed{D}l = y_e e\varphi, \quad i\slashed{D}e = y_e^\dagger\varphi^\dagger l, \quad i\slashed{D}Q = y_u u\tilde{\varphi} + y_d d\varphi, \quad i\slashed{D}_\mu = y_u^\dagger\tilde{\varphi}^\dagger Q, \quad i\slashed{D}d = y_d^\dagger\varphi^\dagger Q. \tag{2.54}$$

$\underline{\psi^2 D^3}$: By reordering derivatives, we can use the equations of motion to reduce these operators to operators of the class $\psi^2\varphi D^2$.

$\underline{\psi^2\varphi D^2}$: All operators in this class can be reduced to (up to total derivatives) operators in the single fermionic current classes $\psi^2\phi^3$ and $\psi^2 X\varphi$ or four-fermion operators ψ^4, plus operators that vanish by the equations of motion.

$(\bar{L}L)(\bar{L}L)$		$(\bar{R}R)(\bar{R}R)$		$(\bar{L}L)(\bar{R}R)$	
$\mathcal{O}_{ll}$	$(\bar{L}\gamma_\mu L)(\bar{L}\gamma^\mu L)$	$\mathcal{O}_{ee}$	$(\bar{e}\gamma_\mu e)(\bar{e}\gamma^\mu e)$	$\mathcal{O}_{le}$	$(\bar{L}\gamma_\mu L)(\bar{e}\gamma^\mu e)$
$\mathcal{O}_{qq}^{(1)}$	$(\bar{Q}\gamma_\mu Q)(\bar{Q}\gamma^\mu Q)$	$\mathcal{O}_{uu}$	$(\bar{u}\gamma_\mu u)(\bar{u}\gamma^\mu u)$	$\mathcal{O}_{Lu}$	$(\bar{L}\gamma_\mu L)(\bar{u}\gamma^\mu u)$
$\mathcal{O}_{qq}^{(3)}$	$(\bar{Q}\gamma_\mu \tau^I Q)(\bar{Q}\gamma^\mu \tau^I Q)$	$\mathcal{O}_{dd}$	$(\bar{d}\gamma_\mu d)(\bar{d}\gamma^\mu d)$	$\mathcal{O}_{Ld}$	$(\bar{L}\gamma_\mu L)(\bar{d}\gamma^\mu d)$
$\mathcal{O}_{Lq}^{(1)}$	$(\bar{L}\gamma_\mu L)(\bar{Q}\gamma^\mu Q)$	$\mathcal{O}_{eu}$	$(\bar{e}\gamma_\mu e)(\bar{u}\gamma^\mu u)$	$\mathcal{O}_{qe}$	$(\bar{Q}\gamma_\mu Q)(\bar{e}\gamma^\mu e)$
$\mathcal{O}_{Lq}^{(3)}$	$(\bar{L}\gamma_\mu \tau^I L)(\bar{Q}\gamma^\mu \tau^I Q)$	$\mathcal{O}_{ed}$	$(\bar{e}\gamma_\mu e)(\bar{d}\gamma^\mu d)$	$\mathcal{O}_{qu}^{(1)}$	$(\bar{Q}\gamma_\mu Q)(\bar{u}\gamma^\mu u)$
		$\mathcal{O}_{ud}^{(1)}$	$(\bar{u}\gamma_\mu u)(\bar{d}\gamma^\mu d)$	$\mathcal{O}_{qu}^{(8)}$	$(\bar{Q}\gamma_\mu T^A Q)(\bar{u}\gamma^\mu T^A u)$
		$\mathcal{O}_{ud}^{(8)}$	$(\bar{u}\gamma_\mu T^A u)(\bar{d}\gamma^\mu T^A d)$	$\mathcal{O}_{qd}^{(1)}$	$(\bar{Q}\gamma_\mu Q)(\bar{d}\gamma^\mu d)$
				$\mathcal{O}_{qd}^{(8)}$	$(\bar{Q}\gamma_\mu T^A Q)(\bar{d}\gamma^\mu T^A d)$

$(\bar{L}R)(\bar{R}L)$ and $(\bar{L}R)(\bar{L}R)$		B-violating	
$\mathcal{O}_{ledq}$	$(\bar{L}e)(\bar{d}Q)$	$\mathcal{O}_{duq}$	$\epsilon^{\alpha\beta\gamma}\epsilon_{jk}[(d^\alpha)^T C u^\beta][(Q^{\gamma j})^T C L^k]$
$\mathcal{O}_{quqd}^{(1)}$	$(\bar{Q}u)\epsilon_{jk}(\bar{Q}d)$	$\mathcal{O}_{qqu}$	$\epsilon^{\alpha\beta\gamma}\epsilon_{jk}[(Q^{\alpha j})^T C Q^{\beta k}][(u^\gamma)^T C e]$
$\mathcal{O}_{quqd}^{(8)}$	$(\bar{Q}T^A u)\epsilon_{jk}(\bar{Q}T^A d)$	$\mathcal{O}_{qqq}^{(1)}$	$\epsilon^{\alpha\beta\gamma}\epsilon_{jk}\epsilon_{mn}[(Q^{\alpha j})^T C Q^{\beta k}][(Q^{\gamma m})^T C L^n]$
$\mathcal{O}_{lequ}^{(1)}$	$(\bar{L}^j e)\epsilon_{jk}(\bar{Q}^k u)$	$\mathcal{O}_{qqq}^{(3)}$	$\epsilon^{\alpha\beta\gamma}(\tau^I \epsilon)_{jk}(\tau^I \epsilon)_{mn}[(Q^{\alpha j})^T C Q^{\beta k}][(Q^{\gamma m})^T C L^n]$
$\mathcal{O}_{lequ}^{(3)}$	$(\bar{L}^j \sigma_{\mu\nu} e)\epsilon_{jk}(\bar{Q}^k \sigma^{\mu\nu} u)$	$\mathcal{O}_{duu}$	$\epsilon^{\alpha\beta\gamma}[(d^\alpha)^T C u^\beta][(u^\gamma)^T C e]$

Table 2.2: *The non-redundant four-fermion $D = 6$ operators in the 'Warsaw basis'. For readability we do not explicitly display the fermion generation indices, but where relevant they are denoted by an extra superscript $\mathcal{O}^{prst} \sim \bar{\psi}_p \psi_r \bar{\psi}_s \psi_t$. For example $\mathcal{O}_{ll}^{1231} = (\bar{e}\gamma_\mu \mu)(\bar{\tau}\gamma^\mu e)$.*

$\underline{\psi^2 X D}$: Using the equations of motion for the gauge field and the Bianchi identities, one finds that all operators in this class can be reduced to operators in the classes $\psi^2 X \varphi$ and $\psi^2 \phi^2 D$, and four-fermion operators ψ^4, plus total derivatives.

The remaining non-redundant operators of the single-fermionic current operator classes are thus all in the $\psi^2 \varphi^3$, $\psi^2 X \varphi$ and $\psi^2 \varphi^2 D$ subclasses. They are displayed in columns 3, 5 and 6 of Tab. 2.1.

Four-fermion operators: Although all operators in this class are all of the simple form ψ^4, they constitute by far the most numerous, though they can be straightforwardly classified. Noting that they are generically constructed out of left-handed fields L and right-handed fields R, they can be constructed out of products of hypercharge zero currents $(\bar{L}L)(\bar{L}L)$, $(\bar{R}R)(\bar{R}R)$ and $(\bar{L}L)(\bar{R}R)$, and a few others of the form $(\bar{L}R)(\bar{R}L)$ and $(\bar{L}R)(\bar{L}R)$, as well as four baryon-number violating operators. Though the equations of

motion cannot be used to whittle down this operator set, the Fierz identity,

$$(\bar{\psi}_L \gamma_\mu \psi_L)(\bar{\chi}_L \gamma_\mu \chi_L) = (\bar{\psi}_L \gamma_\mu \chi_L)(\bar{\chi}_L \gamma_\mu \psi_L) \tag{2.55}$$

as well as the identity for the SU(N) generators

$$T^A_{\mu\nu} T^A_{\alpha\beta} = \frac{1}{2}\delta_{\mu\alpha}\delta_{\nu\beta} - \frac{1}{2N}\delta_{\mu\nu}\delta_{\alpha\beta}, \tag{2.56}$$

can be used. The complete non-redundant set of four-fermion operators are shown in Tab. 2.2.

The 64 operators of those tables complete what is commonly referred to as the 'Warsaw basis' of the $D = 6$ Standard Model effective theory. This is the operator basis used throughout this thesis. Other bases for the $D = 6$ operator set are also commonly used [208–211], and it is merely an exercise in linear algebra to translate between them[¶¶]. Excluding the five B-violating operators, whose effects must be strongly suppressed to respect proton decay bounds, we have 59 independent operators. In fact, one can relax the flavour assumptions and allow all possible flavour combinations to be an independent operator. This increases the operator set to 2499 operators. In order to make an analysis tractable, it is typically assumed that the operators obey minimal flavour violation, so that 59 B-conserving operators form a complete set.

The operator set at $D = 7$ and $D = 8$ has also been computed, and there now exist tools for computing the operator set to arbitrarily high dimension, though not all the redundancies are automatically eliminated. The $D = 7$ operator set all violate lepton number, so are typically not interesting for LHC energies, where lepton number conservation has been demonstrated to an extremely high degree. There are 993 *structures* at $D = 8$ [213], even assuming minimal flavour violation (not all of these correspond to a unique operator, however). In order to avoid the proliferation of large numbers of operators, for phenomenological purposes one typically cuts off the expansion at $D = 6$. Since higher-order terms will be proportional to higher-powers of Λ, then provided there is a large enough separation between the low-energy theory and the cutoff, then this truncation is allowed. The $D = 6$ truncation is typically referred to as the Standard Model Effective Field Theory (SMEFT).

2.6 The top quark sector of the Standard Model effective theory

In order to access the sector of the SMEFT that is relevant for top quark physics at hadron colliders, it is necessary to compute the Feynman rules of the operators in Tables 2.1 and 2.2, and calculate which of them lead to modifications of the processes and observables

---

[¶¶]In fact there are tools which automate this procedure entirely [212].

listed in the first chapter. We will take each of these processes in turn, but first we make some general comments about the modifications of a collider observable due to a $D = 6$ operator.

We will focus first on cross-sections, though the same arguments will apply to decay observables as well. In general, the Lorentz-invariant matrix element $\mathcal{M}$ is related to the differential cross-section in some observable X by

$$\frac{\mathrm{d}\sigma}{\mathrm{d}X} = \int \mathrm{d}\Pi_{\mathrm{LIPS}}\delta^{(4)}(X - X')|\mathcal{M}|^2, \tag{2.57}$$

where the Lorentz-invariant phase space element $\mathrm{d}\Pi_{\mathrm{LIPS}} \sim \mathrm{d}X'$. X may be a one particle inclusive quantity such as the p_T of one of the final state particles, or N-particle inclusive, such as the final state invariant mass. In the presence of $D = 6$ operators, the matrix element is modified to $\mathcal{M}_{\mathrm{full}} = \mathcal{M}_{\mathrm{SM}} + \mathcal{M}_{\mathrm{D6}}$, so that the expression for the squared matrix element is

$$|\mathcal{M}_{\mathrm{full}}|^2 = |\mathcal{M}_{\mathrm{SM}}|^2 + 2\Re\mathcal{M}_{\mathrm{SM}}^*\mathcal{M}_{\mathrm{D6}} + |\mathcal{M}_{\mathrm{D6}}|^2. \tag{2.58}$$

The linear term $2\Re\mathcal{M}_{\mathrm{SM}}^*\mathcal{M}_{\mathrm{D6}}$ is proportional to $1/\Lambda^2$, and is generated by interference between Standard Model and new physics amplitudes, while the quadratic term is generated solely by new physics contributions, and is proportional to $1/\Lambda^4$. Since there is no dependence on the matrix element in the final state phase space, we can schematically write the cross-section as

$$\mathrm{d}\sigma_{\mathrm{full}} \sim \mathrm{d}\sigma_{\mathrm{SM}} + c_i\mathrm{d}\sigma_{\mathrm{D6}} + c_i^2\mathrm{d}\sigma_{\mathrm{D6}^2}, \tag{2.59}$$

where we have displayed the $D = 6$ Wilson coefficients explicitly.

Since the matrix element receives contributions from terms at different orders in Λ, one might worry that the corresponding cross-section is not properly defined in Λ. Namely, the quadratic terms $\mathcal{M}_{\mathrm{D6}}$ and the interference between $D = 8$ operators and the Standard Model $\sim 2\Re\mathcal{M}_{\mathrm{SM}}^*\mathcal{M}_{\mathrm{D8}}$ are formally both $\mathcal{O}(1/\Lambda^4)$, however we only consider the former and not the latter. This issue ties in with a broader discussion of the validity of the EFT description of an observable, and we will return to it throughout this thesis. For now it suffices to assume that the higher-order interference terms can be neglected.

2.6.1 Top pair production

As in the case of the Standard Model, we can (at leading-order in α_s) split up top-quark pair production into the gg and $q\bar{q}$ channels. For the former, any new physics which couples to the top quark directly will modify the top-gluon vertex. The only operator that does this directly is the so-called chromomagnetic moment operator $\mathcal{O}_{uG}^{33} \equiv \mathcal{O}_{tG}$, where the
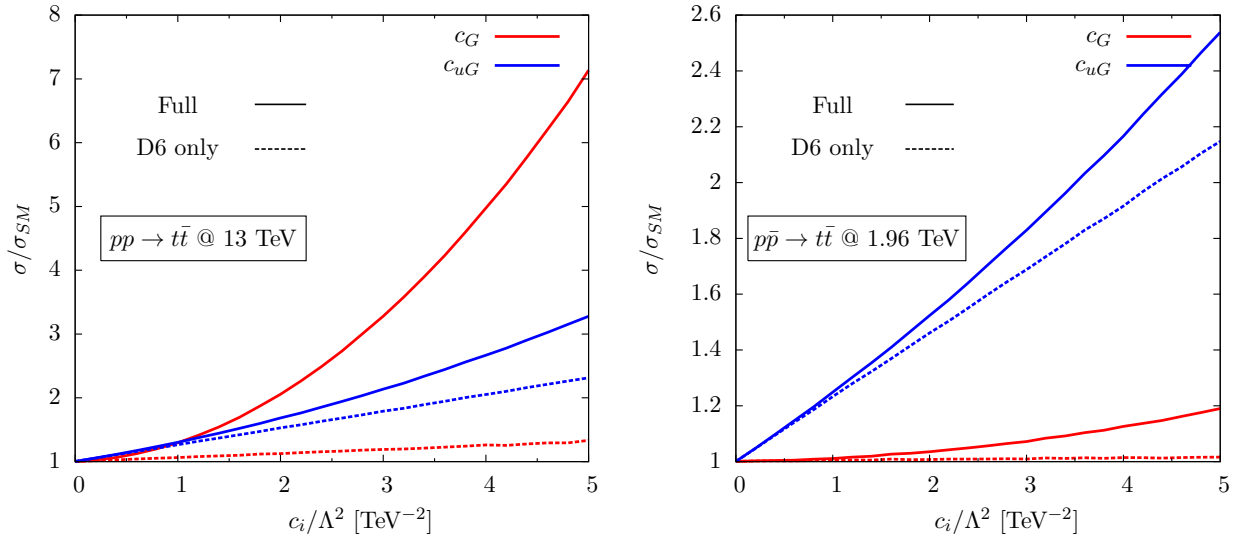
Figure 2.8: *Ratio of the full cross-section for $t\bar{t}$ production at the 13 TeV LHC (left) and Tevatron (right) for the operators $\mathcal{O}_G$ and $\mathcal{O}_{uG}^{33}$ to the Standard model prediction, as a function of the operator Wilson coefficient. The dashed lines show the effects of the interference term only, the solid lines show the interference and quadratic terms.*

superscript denotes the generation index explicitly. Its interference with the SM $gg \to t\bar{t}$ amplitude gives the term [214]***

$$2\Re\mathcal{M}_{uG}^{33*}\mathcal{M}_{\text{SM}} = \frac{g_s^3}{2\sqrt{2}}\frac{vm_t c_{uG}^{33}}{\Lambda^2}\left(\frac{1}{6\tau_1\tau_2} - \frac{3}{8}\right), \tag{2.60}$$

where $\tau_{1,2}$ are functions of the Mandelstam invariants $\tau_1 = (m_t^2 - t)/s$, $\tau_2 = (m_t^2 - u)/s$ and $\rho = 4m_t^2/s$ is the threshold variable.

In fact, this is the only operator that modifies the $gg \to t\bar{t}$ production cross-section by directly coupling to the top. We can indirectly modify the cross-section, however, by modifying the triple-gluon vertex in the initial state with the operator $\mathcal{O}_G$. This leads to the interference matrix element [215, 216]

$$2\Re\mathcal{M}_G^*\mathcal{M}_{\text{SM}} = \frac{9}{8}\frac{c_G g_s^3}{\Lambda^2}\frac{m_t^2(\tau_1 - \tau_2)^2}{\tau_1\tau_2}. \tag{2.61}$$

The partonic differential cross-sections are folded with the incoming parton densities to give the proton-(anti)proton cross-sections. To get an idea of the strength of the operators, we plot the ratio of the full cross-section to the SM only value, broken up into the interference and quadratic pieces, as a function of the dimensional Wilson coefficient $\tilde{c}_i = c_i/\Lambda^2$ for

---

***We want to focus on the interference terms, so we do not show the quadratic terms explicitly, though we will compute their numerical contributions to observables.
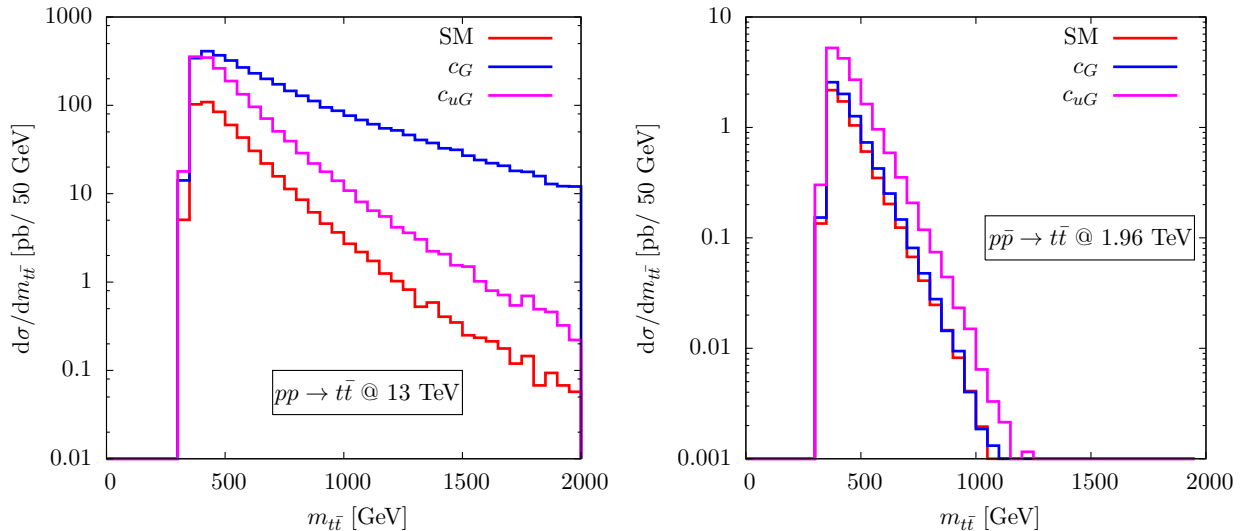
Figure 2.9: *Top quark pair invariant mass distributions at the 13 TeV LHC (left) and the Tevatron (right). The red curves are for the SM only, while the blue and magenta curves are for (respectively) the coefficients c_G/Λ^2 and c_{uG}^{33}/Λ^2 set to a value of 5 TeV $^{-2}$.*

$\mathcal{O}_{uG}^{33}$ and $\mathcal{O}_G$. The results are shown in Fig. 2.8, for both the 13 TeV LHC and Tevatron.

We see that the effects of the operator $\mathcal{O}_G$ come almost entirely from its quadratic term, and its interference is very small, therefore in a $D = 6$ framework, constraints on its coefficient should be taken with caution. It is also clear that $\mathcal{O}_G$ is much stronger at the LHC than at the Tevatron, due to the much higher gluon densities in the proton beam at LHC energies. For the chromomagnetic moment $\mathcal{O}_{uG}^{33}$, the interference term dominates up to very large values of its Wilson coefficient. Its effects are also typically much stronger at the Tevatron, because its larger contribution is from the $q\bar{q}$ channel, which dominates here since both quark and antiquark are valence and dominate over gluons at the typical x values probed at the Tevatron.

We can also show the effects of these two operators on kinematic distributions. In Fig. 2.9 we plot the invariant mass distributions at a point in the parameter space for each operator, namely $c_i/\Lambda^2 = 5$ TeV $^{-2}$, again at the 13 TeV LHC and Tevatron. We see again that in the case of the LHC, the operator $\mathcal{O}_G$ has the much stronger effect, whereas the operator $\mathcal{O}_{uG}^{33}$ dominates at the Tevatron. The operator $\mathcal{O}_{uG}^{33}$, in both cases, modifies the SM distribution by an overall normalisation factor, whereas the operator $\mathcal{O}_G$ has a much stronger effect in the tail of the distribution at the LHC, corresponding to the region where the gluon pdf becomes increasingly dominant.

For the $q\bar{q}$ channel, the situation is slightly more complicated. In addition to the chromomagnetic operator $\mathcal{O}_{uG}^{33}$ already mentioned, there is also a contribution from various four-fermion operators listed in Tab. 2.2. Though there are many individual operators that
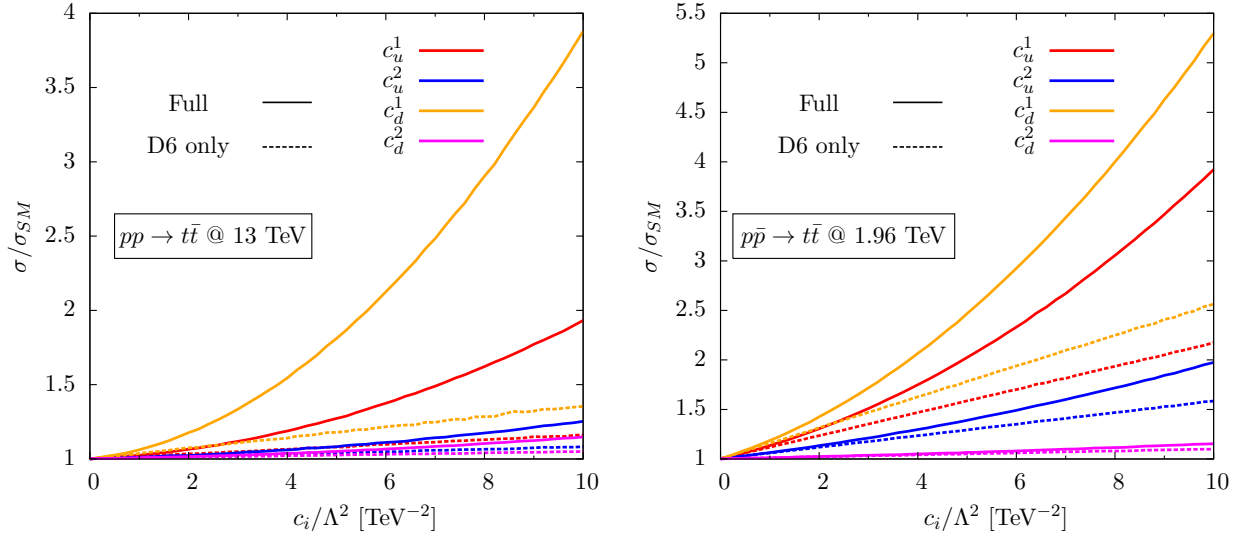
Figure 2.10: *Ratio of the full cross-section for $t\bar{t}$ production at the 13 TeV LHC (left) and Tevatron (right) for the four-fermion operators contributing to top pair production to the Standard model prediction, as a function of the operator Wilson coefficient. The dashed lines show the effects of the interference term only, the solid lines show the interference and quadratic terms.*

contribute, at the level of observables their effects factorise into only four unique linear combinations of 4-quark operators [217].

$$
\begin{aligned}
c_u^1 &= c_{qq}^{1,1331} + c_{uu}^{1331} + c_{qq}^{3,1331} \\
c_u^2 &= c_{qu}^{8,1133} + c_{qu}^{8,3311} \\
c_d^1 &= 4c_{qq}^{3,1133} + c_{ud}^{8,3311} \\
c_d^2 &= c_{qu}^{8,1133} + c_{qd}^{8,3311}.
\end{aligned}
\tag{2.62}
$$

The interference matrix element for these operators is given by

$$
2\Re\mathcal{M}_{4q}^*\mathcal{M}_{\mathrm{SM}} = \frac{g_s^2}{9\pi^2}\frac{s}{\Lambda^2}\left[\frac{1}{4}\left(c_{u,d}^1 - c_{u,d}^2\right)(\tau_1 - \tau_2) + \frac{1}{4}\left(c_{u,d}^1 + c_{u,d}^2\right)\left(\tau_1^2 + \tau_2^2 + \frac{\rho}{2}\right)\right]. \tag{2.63}
$$

The ratio of the full $t\bar{t}$ cross-section (summed over the u and d production channels), including these operators to the SM estimate, at the LHC and Tevatron, are shown in Fig. 2.10. We see that for larger values of the Wilson coefficient, the squared term dominates contributions to the cross-section, and that the $c_{u,d}^1$ type operators typically have a much larger effect on the cross-section, at both the interference and squared level.

2.6.2 Single top production

For single top production, we can split up the processes into s and t channel production, as discussed in the previous chapter. However, since the diagrams are related by crossing symmetry, the same operator set contributes to both. The interfering operators which have numerical significance are $\mathcal{O}_{uW}^{33}$ $(\equiv \mathcal{O}_{tW})$, $\mathcal{O}_{\varphi q}^{(3)}$, and the linear combination of four-fermion operators

$$\mathcal{O}_t \equiv \mathcal{O}_{qq}^{3,1133} + \frac{1}{6}\left(\mathcal{O}_{qq}^{1,1331} + \mathcal{O}_{qq}^{3,1331}\right). \tag{2.64}$$

For t-channel production, the interference term generated by these three operators can be more conveniently expressed directly in terms of the Mandelstam invariants and, for the $ub \to dt$ subprocess, takes the form

$$2\Re\mathcal{M}_{D6}^{*}\mathcal{M}_{\text{SM}} = \frac{c_{\varphi q}^{(3)}}{\Lambda^2}\frac{V_{tb}|V_{ud}|^2 g^2 v^2 s(s-m_t^2)}{4(t-m_W^2)^2} - \frac{c_{tW}}{\Lambda^2}\frac{\sqrt{2}V_{tb}|V_{ud}|^2 m_t m_W st}{(t-m_W^2)^2}$$
$$+ \frac{c_t}{\Lambda^2}\frac{9V_{tb}V_{ud}g^2 s(s-m_t^2)}{8(t-m_W^2)}. \tag{2.65}$$

The corresponding expression for the $\bar{d}b \to \bar{u}t$ subprocess is obtained by substituting $s \to u$ in each term. For s-channel production, the expression can be obtained by substituting $s \to u, t \to s$, making use of crossing symmetry.

The interference of $\mathcal{O}_{\varphi q}^{(3)}$ has the same kinematic dependence as the SM contribution, in fact it amounts to rescaling the CKM element $V_{tb}^2 \to V_{tb}^2 + 2c_{\varphi q}^{(3)}V_{tb}/\Lambda^2$. Therefore it does not affect the shapes of distributions and merely rescales the overall cross-section. The other two operators, in addition to modifying the overall cross-section, modify the shapes of distributions, and so stronger bounds should be expected on their coefficients.

The relative contributions of each of these operators will serve as a rough guide for how strongly they can be constrained when their coefficients are fit to relevant measurements. The next chapter will discuss in detail such a fit. The same arguments can be made to calculate the effects of $D = 6$ operators on the other top-related processes discussed in Chapter 1, such as associated production, charge asymmetries and decay observables. We will postpone such a discussion until the next chapter.

2.7 Summary

Despite the many successes of the Standard Model over the last forty years, there are several reasons to believe that it is only a stepping stone to a more fundamental theory of Nature, due to considerations such as the hierarchy problem, vacuum stability, and gauge coupling unification. Given its large Yukawa coupling, and as the only SM fermion with an electroweak scale mass, the top quark typically plays a special role in most of these

scenarios, such as supersymmetry, extra dimensions and broken global symmetries such as the little Higgs family of models. These extensions generically predict direct and/or indirect modifications to top quark collider observables, such as new resonances decaying to top pairs, enhanced top quark production cross-sections, new top decay modes and modified decay distributions.

A generic way to parametrise the potential effects of heavy new physics on low energy observables is to regard the Standard Model as the leading order piece of an effective theory, where the non-SM interactions are encoded in higher-dimensional ($D > 4$) operators. We formulated the $D = 6$ extension of the Standard Model, and discussed the sector of this EFT that can potentially impact top quark observables at hadron colliders. In the next chapter we will perform a global fit of the Standard Model EFT to Tevatron and LHC Run I data, and will discuss the implications of the results of this fit for constraints on new physics.

3 A global fit of top quark effective theory to data

3.1 Introduction

In chapter 1, we discussed the role of the top quark in the Standard Model of particle physics, and the unique properties of its production and decay mechanisms that can be measured with precision at hadron colliders. In chapter 2, we showed that the top quark also plays a special role in scenarios of physics Beyond the Standard Model, and outlined the formulation of the Standard Model as the leading part of an effective theory. We then briefly touched upon the sector of this effective theory that could be probed with top quark measurements at hadron colliders.

The motivations for this formulation are manifold: Firstly, with the LHC Run II well underway, the main take-home message is that, apart from a few scattered anomalies, all measurements are in agreement with Standard Model predictions. This implies that, if new heavy degrees of freedom exist at all, they are decoupled [218, 219] from the electroweak scale (either there is a large mass gap or very weak coupling between the SM and new physics sectors), in which case they will necessarily integrate out into higher-dimensional operators [195–197, 220] in the low energy limit*. Secondly, faced with the large number of hypothesised new physics scenarios and the frequent degeneracy in their experimental signatures, it is prudent to describe deviations from the Standard Model in as *model-independent* a way as possible. The differences in early inclusive Higgs production cross-section measurements from their SM values, for instance, are often described by 'signal strength' ratios. Likewise, electroweak observables are also phrased in the language of anomalous couplings.

From a phenomenological perspective, the EFT description is nothing more than another model-independent way of asking 'which self-consistent Lagrangian best describes the data?', but it has the advantage over other approaches such as signal strengths in that it can also accommodate differential observables and angular observables, because the higher-dimensional operators lead to new vertex structures which can impact event kinematics. They also have the advantage over 'form-factors' in that they manifestly preserve the full $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry, and so can be more straightforwardly linked to concrete ultraviolet completions.

This merits have not gone unnoticed, and EFT techniques have received much attention in interpreting available Higgs results [222–239], although this area is still in its infancy, and the corresponding bounds (and thus the conclusions that one can draw about extended Higgs sectors) are limited by low statistics on the experimental side. Top quark physics,

---

*Current collider measurements, however, cannot rule out the existence of light degrees of freedom, see e.g. Ref. [221].

by contrast, has entered a precision era, and data from the LHC and Tevatron are far more abundant. As already discussed, the top quark plays a special role in most scenarios of Beyond the Standard Model physics, motivating scrutiny of its phenomenology. Furthermore, the top sector is strongly coupled to Higgs physics owing to the large top quark Yukawa coupling, and so represents a complementary window into physics at the electroweak scale. Thus, it is timely to review the constraints on new top interactions through a global fit of all dimension-six operators relevant to top production and decay at hadron colliders. This is the subject of this chapter.

There have been several studies of the potential for uncovering new physics effects in the top quark sector at the LHC and Tevatron, phrased in model-independent language, either through anomalous couplings [204,240–254] or higher-dimensional operators [214,217,255–259]. Though there is a one-to-one correspondence between these two approaches (for the reasons discussed below) the latter is the approach taken in this analysis. Other studies have also set limits on top dimension-six operators, but by considering different physics, such as precision electroweak data [260], or flavour-changing neutral currents [261,262].

The chapter is structured as follows. In Section 3.2 we review the higher-dimensional operators relevant for top quark physics and in Section 3.3 we review the experimental measurements entering our fit, as well as the limit-setting procedure we adopt. In Section 3.4 we present our constraints, and discuss the complementarity of LHC and Tevatron analyses, and the improvements obtained from adding differential distributions as well as inclusive rates. In Section 3.5 we discuss issues relating to the validity of the EFT framework and how our constraints look in the context of specific new physics models. Finally, in Section 3.6 we discuss our results and conclude.

3.2 Higher-dimensional operators

In order to keep this chapter self-contained, in this section we briefly revisit the operators relevant for top observables at hadron colliders. As discussed in chapter 2, the leading contributions to $\mathcal{L}_{\text{eff}}$ at collider energies enter at dimension $D = 6$

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{SM}} + \frac{1}{\Lambda^2} \sum_i c_i \mathcal{O}_i(G_\mu^A, W_\mu^I, B_\mu, \varphi, Q_L, u_R, d_R, L_L, e_R) + \mathcal{O}(\Lambda^{-4}). \qquad (3.1)$$

$\mathcal{O}_i$ are $D = 6$ operators made up of SM fields, and c_i are dimensionless Wilson coefficients. At dimension-six, assuming minimal flavour violation and Baryon number conservation, there are 59 independent operators. Clearly, allowing 59 free parameters to float in a global fit is intractable. Fortunately, for any given class of observables, only a smaller subset is relevant. In top physics, for the observables we consider, we have the following

effective operators, expressed in the so-called 'Warsaw basis' of Ref. [199][†]

$$
\begin{aligned}
&\mathcal{O}_{qq}^{(1)} = (\bar{Q}\gamma_\mu Q)(\bar{Q}\gamma^\mu Q) &&\mathcal{O}_{uW} = (\bar{Q}\sigma^{\mu\nu}\tau^I u)\tilde{\varphi}W_{\mu\nu}^I &&\mathcal{O}_{\varphi q}^{(3)} = i(\varphi^\dagger \overset{\leftrightarrow}{D}{}_\mu^I \varphi)(\bar{Q}\gamma^\mu \tau^I Q) \\
&\mathcal{O}_{qq}^{(3)} = (\bar{Q}\gamma_\mu \tau^I Q)(\bar{Q}\gamma^\mu \tau^I Q) &&\mathcal{O}_{uG} = (\bar{Q}\sigma^{\mu\nu}T^A u)\tilde{\varphi}G_{\mu\nu}^A &&\mathcal{O}_{\varphi q}^{(1)} = i(\varphi^\dagger \overset{\leftrightarrow}{D}{}_\mu \varphi)(\bar{Q}\gamma^\mu Q) \\
&\mathcal{O}_{uu} = (\bar{u}\gamma_\mu u)(\bar{u}\gamma^\mu u) &&\mathcal{O}_G = f_{ABC}G_\mu^{A\nu}G_\nu^{B\lambda}G_\lambda^{C\mu} &&\mathcal{O}_{uB} = (\bar{Q}\sigma^{\mu\nu}u)\tilde{\varphi}B_{\mu\nu} \\
&\mathcal{O}_{qu}^{(8)} = (\bar{Q}\gamma_\mu T^A Q)(\bar{u}\gamma^\mu T^A u) &&\mathcal{O}_{\tilde{G}} = f_{ABC}\tilde{G}_\mu^{A\nu}G_\nu^{B\lambda}G_\lambda^{C\mu} &&\mathcal{O}_{\varphi u} = (\varphi^\dagger i\overset{\leftrightarrow}{D}{}_\mu \varphi)(\bar{u}\gamma^\mu u) \\
&\mathcal{O}_{qd}^{(8)} = (\bar{Q}\gamma_\mu T^A Q)(\bar{d}\gamma^\mu T^A d) &&\mathcal{O}_{\varphi G} = (\varphi^\dagger \varphi)G_{\mu\nu}^A G^{A\mu\nu} &&\mathcal{O}_{\varphi\tilde{G}} = (\varphi^\dagger \varphi)\tilde{G}_{\mu\nu}^A G^{A\mu\nu} \\
&\mathcal{O}_{ud}^{(8)} = (\bar{u}\gamma_\mu T^A u)(\bar{d}\gamma^\mu T^A d)\,.
\end{aligned}
\tag{3.2}
$$

We adopt the same notation as Ref. [199], where $T^A = \frac{1}{2}\lambda^A$ are the $SU(3)$ generators, and τ^I are the Pauli matrices, related to the generators of $SU(2)$ by $S^I = \frac{1}{2}\tau^I$. For the four-quark operators on the left column of Eq. (3.2), we denote a specific flavour combination $(\bar{Q}_i...Q_j)(\bar{Q}_k...Q_l)$ by e.g. $\mathcal{O}_{ijkl}^{4q}$. It should be noted that the operators $\mathcal{O}_{uW}$, $\mathcal{O}_{uG}$ and $\mathcal{O}_{uB}$ are not hermitian and so may have complex coefficients which, along with $\mathcal{O}_{\tilde{G}}$ and $\mathcal{O}_{\varphi\tilde{G}}$, lead to $\mathcal{CP}$-violating effects. These do not contribute to the observables built out of spin-averaged matrix elements that we consider, but they are in principle sensitive to polarimetric information such as spin correlations, and should therefore be treated as independent operators. However, currently available measurements that would be sensitive to these degrees of freedom have been extracted by making model-specific assumptions that preclude their usage in the fit, e.g. by assuming that the tops are produced with either SM-like spin correlation or no spin correlation at all, as in Refs. [266–268]. We will discuss this issue in more detail in the next section. With these caveats, a total of 14 constrainable $\mathcal{CP}$-even dimension-six operators contribute to top quark production and decay at leading order in the SMEFT.

3.3 Methodology

3.3.1 Experimental inputs

The experimental measurements used in the fit [269–306] are included in Table 3.1. All these measurements are quoted in terms of 'parton-level' quantities; that is, top quarks and their direct decay products. Whilst it is possible to include particle-level observables, these are far less abundant and they are beyond the scope of the present study.

The importance of including kinematic distributions is manifest here. For top pair production, for instance, we have a total of 195 measurements, 174 of which come from

---

[†]Given the simplicity of how it captures modifications to SM fermion couplings, this basis is well-suited to top EFT. For basis choices of interest in Higgs physics, see e.g. Refs. [209, 211, 263–265], and Ref. [212] for a tool for translating between them.

Dataset	$\sqrt{s}$ (TeV)	Measurements	arXiv ref.	Dataset	$\sqrt{s}$ (TeV)	Measurements	arXiv ref.
Top pair production							
Total cross-sections:				Differential cross-sections:			
ATLAS	7	lepton+jets	1406.5375	ATLAS	7	$p_T(t), m_{t\bar{t}}, |y_{t\bar{t}}|$	1407.0371
ATLAS	7	dilepton	1202.4892	CDF	1.96	$m_{t\bar{t}}$	0903.2850
ATLAS	7	lepton+tau	1205.3067	CMS	7	$p_T(t), m_{t\bar{t}}, y_t, y_{t\bar{t}}$	1211.2220
ATLAS	7	lepton w/o b jets	1201.1889	CMS	8	$p_T(t), m_{t\bar{t}}, y_t, y_{t\bar{t}}$	1505.04480
ATLAS	7	lepton w/ b jets	1406.5375	D$\emptyset$	1.96	$m_{t\bar{t}}, p_T(t), |y_t|$	1401.5785
ATLAS	7	tau+jets	1211.7205				
ATLAS	7	$t\bar{t}, Z\gamma, WW$	1407.0573	Charge asymmetries:			
ATLAS	8	dilepton	1202.4892	ATLAS	7	A_C (inclusive+$m_{t\bar{t}}, y_{t\bar{t}}$)	1311.6742
CMS	7	all hadronic	1302.0508	CMS	7	A_C (inclusive+$m_{t\bar{t}}, y_{t\bar{t}}$)	1402.3803
CMS	7	dilepton	1208.2761	CDF	1.96	A_{FB} (inclusive+$m_{t\bar{t}}, y_{t\bar{t}}$)	1211.1003
CMS	7	lepton+jets	1212.6682	D$\emptyset$	1.96	A_{FB} (inclusive+$m_{t\bar{t}}, y_{t\bar{t}}$)	1405.0421
CMS	7	lepton+tau	1203.6810				
CMS	7	tau+jets	1301.5755	Top widths:			
CMS	8	dilepton	1312.7582	D$\emptyset$	1.96	Γ_{top}	1308.4050
CDF + D$\emptyset$	1.96	Combined world average	1309.7570	CDF	1.96	Γ_{top}	1201.4156
Single top production				*W*-boson helicity fractions:			
ATLAS	7	t-channel (differential)	1406.7844	ATLAS	7		1205.2484
CDF	1.96	s-channel (total)	1402.0484	CDF	1.96		1211.4523
CMS	7	t-channel (total)	1406.7844	CMS	7		1308.3879
CMS	8	t-channel (total)	1406.7844	D$\emptyset$	1.96		1011.6549
D$\emptyset$	1.96	s-channel (total)	0907.4259				
D$\emptyset$	1.96	t-channel (total)	1105.2788				
Associated production				*Run II data*			
ATLAS	7	$t\bar{t}\gamma$	1502.00586	CMS	13	$t\bar{t}$ (dilepton)	1510.05302
ATLAS	8	$t\bar{t}Z$	1509.05276				
CMS	8	$t\bar{t}Z$	1406.7830				

Table 3.1: *The measurements entering the fit. Details of each are described in the text.*

differential observables. This size of fit is unprecedented in top physics, which underlines the need for a systematic fitting approach, as provided by PROFESSOR [307]. Indeed top pair production cross-sections make up the bulk of measurements that are used in the fit. Single top production cross-sections comprise the next dominant contribution. We also make use of data from charge asymmetries in top pair production, as well as inclusive measurements of top pair production in association with a photon or a Z ($t\bar{t}\gamma$ and $t\bar{t}Z$) and observables relating to top quark decay. We take each of these categories of measurement in turn, discussing which operators are relevant and the constraints obtained on them from data.

3.3.2 Treatment of uncertainties

The uncertainties entering the fit can be classed into three categories:

Experimental uncertainties: We generally have no control over these. In cases where statistical and systematic (and luminosity) errors are recorded separately, we add them in quadrature. Correlations between measurements are also an issue: the unfolding of measured distributions to parton-level introduces some correlation between neighbouring

bins. If estimates of these effects have been provided in the experimental analysis, we use this information in the fit, if they are not, we assume zero correlation. However, we have checked that bin correlations have little effect on our numerical results.

There will also be correlations between apparently separate measurements. The multitude of different top pair production cross-section measurements will clearly be correlated due to overlapping event selection criteria and detector effects, etc. Without a full study of the correlations between different decay channels measured by the same experiment, these effects cannot be completely taken into account, but based on the negligible effects of the bin-by-bin correlations on our numerical results we can expect these effects to be small as well.

Standard Model theoretical uncertainties: These stem from the choice of parton distribution functions (PDFs), as well as neglected higher-order perturbative corrections. As discussed in chapter 1, we model the latter by varying the renormalisation and factorisation scales independently in the range $\mu_0/2 \leq \mu_{\mathrm{R,F}} \leq 2\mu_0$, where we use $\mu_0 = m_t$ as the default scale, and take the envelope as our uncertainty. For the PDF uncertainty, we follow the PDF4LHC recommendation [308] of using CT10 [309], MSTW [96] & NNPDF [310] NLO fits, each with associated scale uncertainties, then taking the full width of the scale+PDF envelope as our uncertainty estimate – i.e. we conservatively assume that scales and parton densities are 100% correlated. Unless otherwise stated, we take the top quark mass to be $m_t = 173.2 \pm 1.0$ GeV. We do not consider electroweak corrections.

Only recently has a lot of progress been made in extending the dimension six-extended SM to higher order in α_s, see Refs. [311–324]. Including these effects is beyond the scope of this work, also because we work to leading order accuracy in the electroweak expansion of the SM. QCD corrections to four fermion operators included via renormalisation group equations are typically of the order of 15%, depending on the resolved phase space [320]. As pointed out in Ref. [325], these effects can be important in electroweak precision data fits.

Interpolation error: A small error relating to the Monte Carlo interpolation (described in more detail in the next section) is included. This is estimated to be 5% at a conservative estimate, as discussed in the following section, and thus subleading compared to the previous two categories.

3.3.3 Fitting procedure

Our fitting procedure, briefly outlined in Ref. [1], uses the PROFESSOR framework. The first step is to construct an N-dimensional hypercube in the space of dimension six couplings,
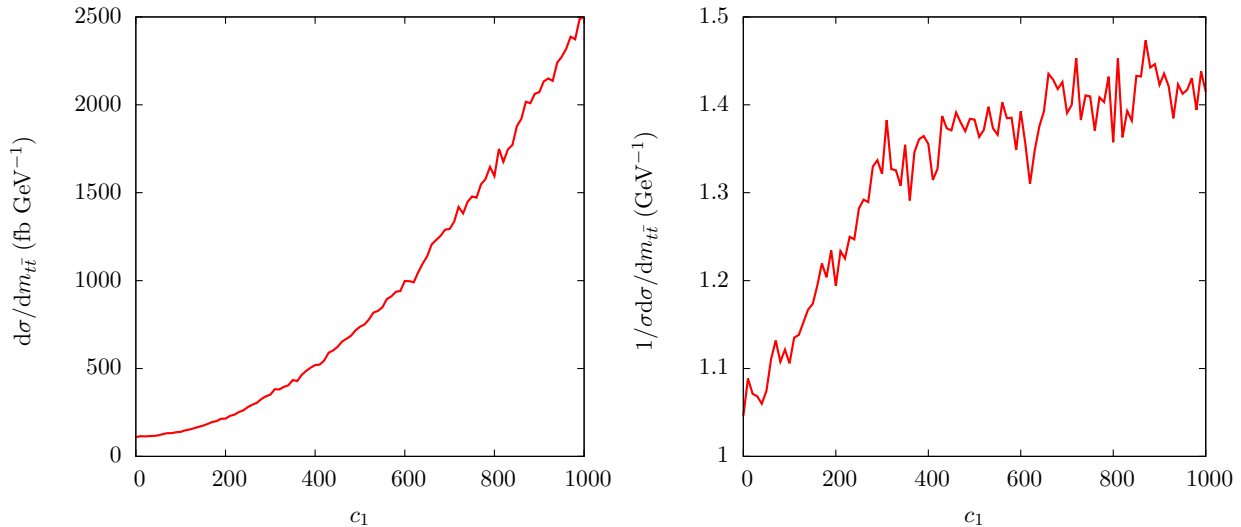
Figure 3.1: *Dependence of a bin in the unnormalised (left) and normalised (right) $t\bar{t}$ invariant mass distribution on one of the dimension-six coefficients considered in this fit.*

compute the observables at each point in the space, and then to fit an *interpolating function* $f(\mathbf{c})$ that parametrises the theory prediction as a function of the Wilson coefficients $\mathbf{c} = \{c_i\}$. This can then be used to rapidly generate theory observables for arbitrary values of the coefficients. Motivated by the dependence of the total cross-section with a Wilson coefficient (also shown in Eq. 2.59):

$$\sigma \sim \sigma_{\mathrm{SM}} + c_i \sigma_{D6} + c_i^2 \sigma_{D6^2} \,, \tag{3.3}$$

the fitting function is chosen to be a second-order or higher polynomial:

$$f_b(\{c_i\}) = \alpha_0^b + \sum_i \beta_i^b c_i + \sum_{i \leq j} \gamma_{i,j}^b c_i c_j + \dots . \tag{3.4}$$

In the absence of systematic uncertainties, each unnormalised observable would exactly follow a second-order polynomial in the coefficients, and higher-order terms capture bin uncertainties which modify this. The polynomial also serves as a useful check that the dimension-six approximation is valid. By comparing Eq. (3.3) with Eq. (3.4), we see that the terms quadratic in c_i are small provided that the coefficients in the interpolating function $\gamma_{i,j}$ are small. This is a more robust way to ensure validity of the dimension-six approximation than to assume a linear fit from the start.

The simple quadratic dependence on the coefficients is not guaranteed to propagate into every observable, however. Many of the differential distributions recorded by the LHC experiments (and used in this fit) are normalised to unity. This has the advantage of dividing out many of the systematic uncertainties entering the extraction of the cross-
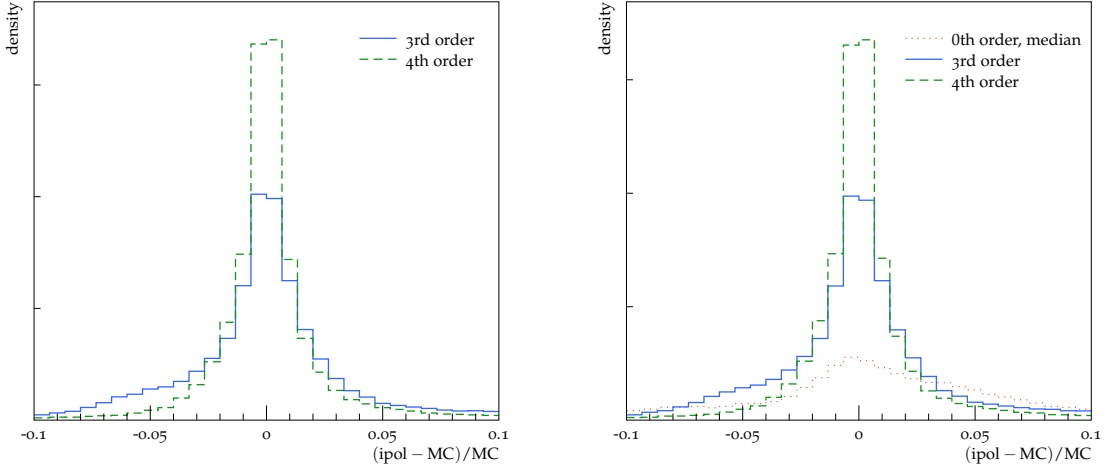
Figure 3.2: *Residuals distributions for interpolated observable values (left) and uncertainties (right), evaluated over all input MC runs and all observables. The 4th order polynomial parameterisation gives the best performance and the vast majority of entries are within 5% of the explicit MC value. The poor performance of a constant uncertainty assumption based on the median input uncertainty is evident – since all three lines have the same normalisation, the majority of residual mismodellings for the median approach are (far) outside the displayed 10% interval.*

section, but has the practical disadvantage from a fitting perspective that the observables follow a much more complicated polynomial dependence on the coefficients. Consider, for example, a bin in the normalised differential cross-section in some observable X, which will schematically take the form

$$\frac{1}{\sigma(c_i)}\frac{d\sigma(c_i)}{dX} \sim \frac{1}{f + gc_i + hc_i^2} \times (f' + g'c_i + h'c_i^2), \tag{3.5}$$

where $\{f, g, h, f', g', h'\}$ are dimensionful functions of kinematic and Standard Model parameters. It is clear to see that this function will not be quadratic in c_i. This is exemplified in figure 3.1, where we show the dependence of a given bin in the top pair invariant mass distribution, both unnormalised and normalised. A clear quadratic dependence is seen in the former, whereas the latter is much more irregular and should be modelled by a higher-order function.

In principle there is no limit on how high a polynomial order we may use. The limiting factor is the increased computation time at each successive order, and the inefficiency of overfitting to statistical noise, which very high order polynomials almost certainly do. In practice, to minimise the interpolation uncertainty, we use up to a 4th order polynomial in Eq. (3.4), depending on the observable of interest. The performance of the interpolation
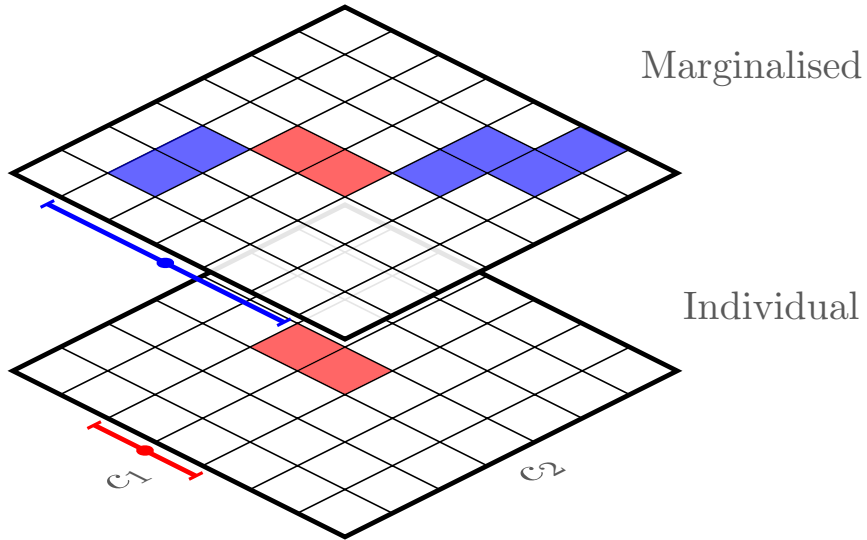
Figure 3.3: *Illustration of the difference between individual and marginalised constraints for a two-parameter fit. The allowed region for c_1 with all other coefficients set to zero (red squares) can be made larger by varying c_1 and c_2 simultaneously and tuning them so that the theory prediction is close to the data (blue squares).*

method is shown in Fig. 3.2, which depicts the fractional deviation of the polynomial fit from the explicit MC points used to constrain it. The central values and the sizes of the modelling uncertainties may both be parameterised with extremely similar performance, with 4th order performing best for both. The width of this residual mismodeling distribution being $\sim 3\%$ for each of the value and error components is the motivation for a total 5% interpolation uncertainty to be included in the goodness of fit of the interpolated MC polynomial $f(\mathbf{c})$ to the experimentally measured value E:

$$\chi^2(\mathbf{c}) = \sum_{\mathcal{O}} \sum_{i,j} \frac{(f_i(\mathbf{c}) - E_i)\rho_{i,j}(f_j(\mathbf{c}) - E_j)}{\Delta_i \Delta_j} \,, \tag{3.6}$$

where we sum over all observables $\mathcal{O}$ and all bins in that observable i. We include the correlation matrix $\rho_{i,j}$ where this is provided by the experiments, otherwise $\rho_{i,j} = \delta_{ij}$. The uncertainty on each bin is given by $\Delta_i = \sqrt{\Delta_{\text{th},i}^2 + \Delta_{\text{exp},i}^2}$, i.e. we treat theory and experimental errors as uncorrelated. The parameterisation of the theory uncertainties is restricted to not become larger than in the training set, to ensure that polynomial blow-up of the uncertainty at the edges of the sampling range cannot produce a spuriously low χ^2 and disrupt the fit.

We hence have constructed a fast parameterisation of model goodness-of-fit as a function of the EFT operator coefficients. This may be used to produce χ^2 maps in slices (where all but one parameter is fixed, typically to zero) or marginalised (where all parameters are
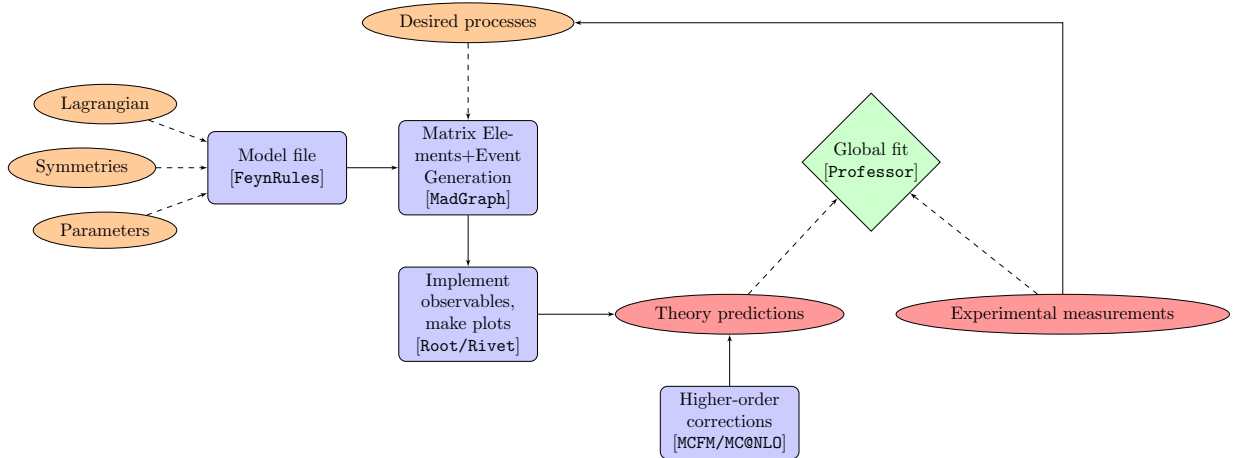
Figure 3.4: *Flowchart illustrating the fitting procedure and relevant software used at each step of the analysis.*

varied simultaneously) projections of the operator space. Since in the marginalised case, a pull of the theory prediction in one direction by one parameter can be compensated by tuning another parameter to pull it back into agreement with the data, the net result is that marginalised confidence intervals are wider than individual 'slices'. A visualisation of this is shown in Fig. 3.3 for a two parameter fit. These projections are transformed to confidence intervals on the coefficients c_i, defined by the regions for which

$$1 - \mathrm{CL} \geq \int_{\chi^2(c_i)}^{\infty} f_k(x)dx \,, \tag{3.7}$$

where typically $\mathrm{CL} \in \{0.68, 0.95, 0.99\}$ and $f_k(x)$ is the χ^2 distribution for k degrees of freedom, which we define as $k = N_{\mathrm{measurements}} - N_{\mathrm{coefficients}}$.

A flowchart of the fitting procedure is shown in Fig. 3.4.

3.4 Results

The entire 59 dimensional operator set of Ref. [199] was implemented in a FeynRules [326] model file, with care taken to ensure consistent redefinitions of the relation between SM input parameters and observables (see appendix A for details). The contributions to parton level cross-sections and decay observables from the above operators were computed using MADGRAPH/MADEVENT [101], making use of the Universal FeynRules Output (UFO) [327] format. We model NLO QCD corrections by including Standard Model K-factors (bin-by-bin for differential observables), where the NLO observables are calculated using MCFM [107], cross-checked with MC@NLO [106, 328]. These K-factors are used for arbitrary values of the Wilson coefficients, thus modelling NLO effects in the pure-SM
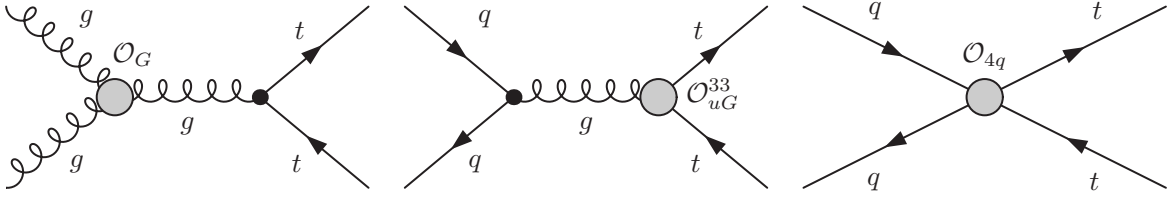
Figure 3.5: *Sample Feynman diagrams for the interference of the leading-order SM amplitudes for top pair production with the operators of Eq. (3.8). $\mathcal{O}_{4q}$ denotes the insertion of any of the four-quark operators.*

contribution only. More specifically, this amounts to performing a simultaneous expansion of each observable in the strong coupling α_s and the (inverse) new physics scale Λ^{-1}, and neglecting terms $\sim \mathcal{O}(\alpha_S \Lambda^{-2})$. For t-channel single top production, all our results are presented in a five-flavour scheme for the incoming pdfs, however we have cross-checked our results against a four-flavour scheme and found good agreement. Our final 95% confidence limits for each coefficient are presented in Fig. 3.18; we discuss them in more detail below.

3.4.1 Top pair production

By far the most abundant source of data in top physics is from the production of top pairs. The $\mathcal{CP}$-even dimension-six operators that interfere with the Standard Model amplitude are

$$
\begin{aligned}
\mathcal{L}_{\text{D6}} \supset & \frac{c_{uG}}{\Lambda^2}(\bar{Q}\sigma^{\mu\nu}T^A u)\tilde{\varphi}G^A_{\mu\nu} + \frac{c_G}{\Lambda^2}f_{ABC}G^{A\nu}_\mu G^{B\lambda}_\nu G^{C\mu}_\lambda + \frac{c_{\varphi G}}{\Lambda^2}(\varphi^\dagger\varphi)G^A_{\mu\nu}G^{A\mu\nu} \\
& + \frac{c^1_{qq}}{\Lambda^2}(\bar{Q}\gamma_\mu Q)(\bar{Q}\gamma^\mu Q) + \frac{c^3_{qq}}{\Lambda^2}(\bar{Q}\gamma_\mu\tau^I Q)(\bar{Q}\gamma^\mu\tau^I Q) + \frac{c_{uu}}{\Lambda^2}(\bar{u}\gamma_\mu u)(\bar{u}\gamma^\mu u) \\
& + \frac{c^8_{qu}}{\Lambda^2}(\bar{Q}\gamma_\mu T^A Q)(\bar{u}\gamma^\mu T^A u) + \frac{c^8_{qd}}{\Lambda^2}(\bar{Q}\gamma_\mu T^A Q)(\bar{d}\gamma^\mu T^A d) + \frac{c^8_{ud}}{\Lambda^2}(\bar{u}\gamma_\mu T^A u)(\bar{d}\gamma^\mu T^A d)\,.
\end{aligned}
$$

$$(3.8)$$

As pointed out in Ref. [1], the operator $\mathcal{O}_{\varphi G}$ cannot be bounded by top pair production alone, since the branching ratio to virtual top pairs for a 125 GeV Higgs is practically zero, therefore we do not consider it here. For a recent constraint from Higgs physics see e.g. Ref. [232, 234, 238, 239]. We further ignore the contribution of the operator $\mathcal{O}^{11}_{uG}$, as this operator is a direct mixing of the left- and right- chiral u quark fields, and so contributes terms proportional to m_u. We also note that the six four-quark operators of Eq. (3.8) interfere with the Standard Model QCD processes $\bar{u}u,\,\bar{d}d \to \bar{t}t$ to produce terms dependent only on the four linear combinations of Wilson Coefficients: $c^{1,2}_{u,d}$, displayed in
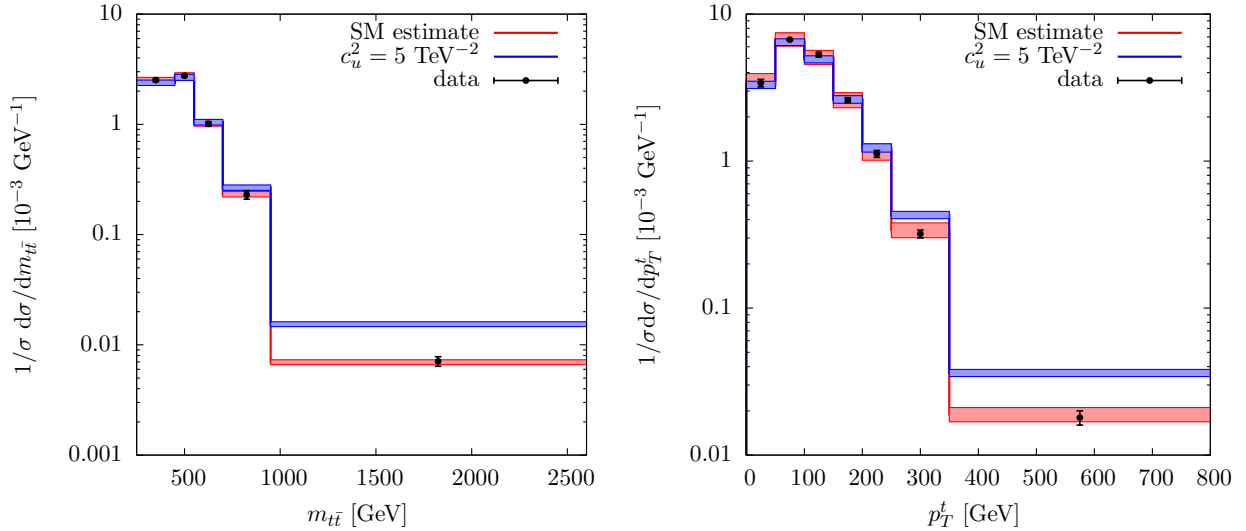
Figure 3.6: *Parton level differential distributions in top pair production, considering SM only (red) and the effects of the four-quark operator $\mathcal{O}_u^2$, showing the enhancement in the tails of the distributions. Data taken from Ref. [289].*

Eq. (2.62).

It is these four that are constrainable in a dimension-six analysis. Finally, we note that the operator $\mathcal{O}_G$, whilst not directly coupling to the top at tree-level, should not be neglected. Since it modifies the triple gluon vertex, and the gg channel contributes $\sim 75\%$ (90%) of the total top pair production cross-section at the 8 (13) TeV LHC, moderate values of its Wilson coefficient can substantially impact total rates, as we already saw in chapter 2. We note, however, that in this special case, the cross section modifications are driven by the squared dimension six terms instead of the linearised interference with the SM. Nonetheless, in the interests of generality, we choose to include this operator in the fit at this stage, noting that bounds on its Wilson coefficient should be interpreted with caution.[‡] Representative Feynman diagrams for the interference of these operators are shown in Fig. 3.5.

The most obvious place to look for the effects of higher-dimensional terms is through the enhancement (or reduction, in the case of destructive interference) of total cross-sections. Important differences between SM and dimension-six terms are lost in this approach, however, since operators can cause deviations in the shape of distributions without substantially impacting event yields. This is highlighted in Fig. 3.6, where we plot our NLO SM estimate for two top pair kinematic distributions, vs. one with a large new physics interference term. Both are consistent with the data in the threshold region, which dominates

---

[‡]We have observed that excluding this operator actually tightens the bounds on the remaining ones, so choosing to keep it is the more conservative option.
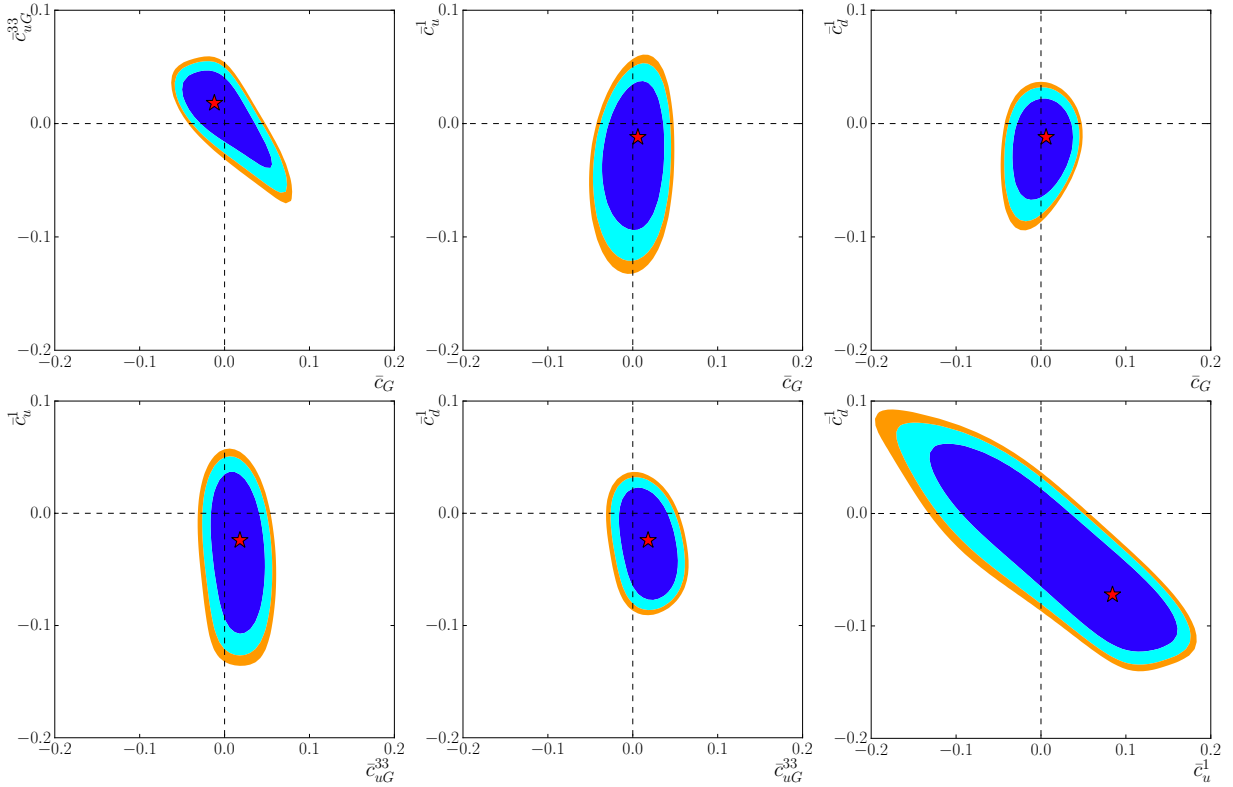
Figure 3.7: *68%, 95% and 99% confidence intervals for selected combinations of operators contributing to top pair production, with all remaining operators set to zero. The star marks the best fit point, indicating good agreement with the Standard Model. Here $\bar{c}_i = c_i v^2 / \Lambda^2$.*

the cross-section, but clear discrimination between SM and dimension-six effects is visible in the high-mass region, which simply originates from the scaling of dimension-six operator effects as s/Λ^2.

Limits on these operators can be obtained in two ways; by setting all other operators to zero, and by marginalising over the other parameters in a global fit. In Fig. 3.7 we plot the allowed 68%, 95% and 99% confidence intervals for various pairs of operators, with all others set to zero, showing correlations between some coefficients. Most of these operators appear uncorrelated, though there is a strong correlation between c_u^1 and c_d^1, due to a relative sign between their interference terms. Given the lack of reported deviations in top quark measurements, it is perhaps unsurprising to see that all Wilson coefficients are consistent with zero within the 95% confidence intervals, and that the SM hypothesis is an excellent description of the data. In Fig. 3.8, the stronger joint constraints on c_G vs c_u^1 obtained from including differential measurements make manifest the importance of utilizing all available cross-section information.

It is also interesting to note the complementarity between measurements from the LHC and Tevatron, as illustrated in Fig. 3.8. It is interesting to see that although Tevatron
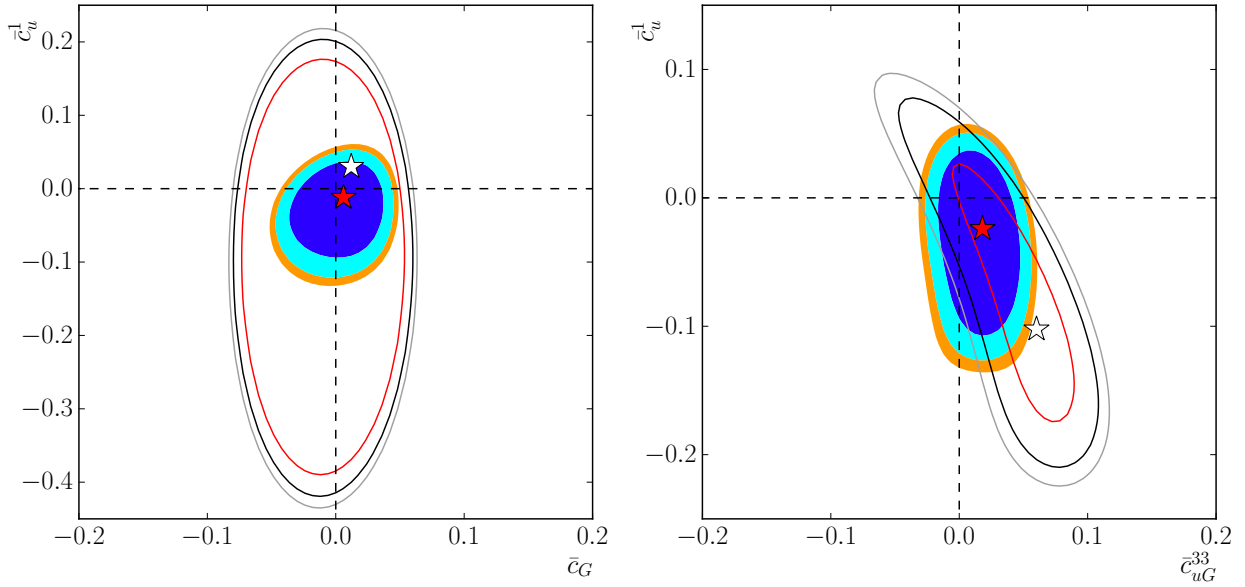
Figure 3.8: *Left: 68%, 95% and 99% confidence intervals on the operators c_G vs. c_u^1, considering differential and total cross-sections (contours, red star), and total cross-sections only (lines, white star). Right: Limits on c_{uG}^{33} vs. c_u^1, considering both Tevatron and LHC data (contours) and Tevatron data only (lines).*

data are naively more sensitive to four-quark operators, after the LHC Run I and early into Run II, the LHC data size and probed energy transfers lead to comparably stronger constraints. In the fit this is highlighted by the simple fact that LHC data comprise more than 80% of the bins, so have a much larger pull. This stresses the importance of collecting large statistics as well as using sensitive discriminating observables.

3.4.2 Single top production

The next most abundant source of top quark data is from single top production. In the fit we consider production in the t and s channels, and omit Wt-associated production. Though measurements of the latter process have been published, they are not suitable for inclusion in a fit involving parton level theory predictions. As discussed in chapter 1, Wt production interferes with top pair production at NLO and beyond in a five-flavour scheme [127, 128, 130], or at LO in a four-flavour one. Its separation from top pair production is then a delicate issue, discussed in detail in Refs. [129, 133, 135, 136]. We thus choose to postpone the inclusion of Wt production to a future study, going beyond parton level.

The operators that could lead to deviations from SM predictions are shown in Eq. (3.9).

$$\mathcal{L}_{\text{D6}} \supset \frac{c_{uW}}{\Lambda^2}(\bar{Q}\sigma^{\mu\nu}\tau^I u)\,\tilde{\varphi}\,W^I_{\mu\nu} + \frac{c^{(3)}_{\varphi q}}{\Lambda^2}i(\varphi^\dagger \overleftrightarrow{D}^I_\mu \varphi)(\bar{Q}\gamma^\mu \tau^I Q)$$
$$+ \frac{c_{\varphi ud}}{\Lambda^2}(\varphi^\dagger \overleftrightarrow{D}_\mu \varphi)(\bar{u}\gamma^\mu d) + \frac{c_{dW}}{\Lambda^2}(\bar{Q}\sigma^{\mu\nu}\tau^I d)\,\tilde{\varphi}\,W^I_{\mu\nu} \qquad (3.9)$$
$$+ \frac{c^3_{qq}}{\Lambda^2}(\bar{Q}\gamma_\mu \tau^I Q)(\bar{Q}\gamma^\mu \tau^I Q) + \frac{c^1_{qq}}{\Lambda^2}(\bar{Q}\gamma_\mu Q)(\bar{Q}\gamma^\mu Q) + \frac{c^1_{qu}}{\Lambda^2}(\bar{Q}\gamma_\mu Q)(\bar{u}\gamma^\mu u)\,.$$

As in top pair production there are several simplifications which reduce this operator set. The right-chiral down quark fields appearing in $\mathcal{O}_{uW}$ and $\mathcal{O}_{\varphi ud}$ cause these operators' interference with the left-chiral SM weak interaction to be proportional to the relevant down-type quark mass. For example, an operator insertion of $\mathcal{O}^{33}_{\varphi ud}$ will always contract with the SM Wtb-vertex to form a term of order $m_b\, m_t\, c^{33}_{\varphi ud}/\Lambda^2$. Since m_b is much less than both $\hat{s}$ and the other dimensionful parameters that appear, v and m_t, we may choose to neglect these operators. By the same rationale we neglect $\mathcal{O}^{(1)}_{qu}$ as its contribution to observables is proportional to m_u. We have further checked numerically that the contribution of these operators is practically negligible. Finally, all contributing four-fermion partonic subprocesses depend only on the linear combination of Wilson Coefficients:

$$c_t = c^{3,1133}_{qq} + \tfrac{1}{6}(c^{1,1331}_{qq} - c^{3,1331}_{qq}). \qquad (3.10)$$

Single top production can thus be characterised by the three dimension-six operators $\mathcal{O}_{uW}$, $\mathcal{O}^{(3)}_{\varphi q}$ and $\mathcal{O}_t$. The correlations among the constraints of these operators are displayed in figure 3.9.

As noted in the introduction to this chapter, several model-independent studies have noted the potential for uncovering new physics in single top production, though these have typically been expressed in terms of anomalous couplings, via the Lagrangian

$$\mathcal{L}_{Wtb} = \frac{g}{\sqrt{2}}\bar{b}\gamma^\mu(V_L P_L + V_R P_R)tW^-_\mu + \frac{g}{\sqrt{2}}\bar{b}\frac{i\sigma^{\mu\nu}q_\nu}{M_W}(g_L P_L + g_R P_R)tW^-_\mu + h.c. \qquad (3.11)$$

where $q = p_t - p_b$. There is a one-to-one mapping between this Lagrangian and those dimension-six operators that modify the Wtb vertex:

$$V_L \to V_{tb} + c^{(3)}_{\varphi q}v^2/\Lambda^2 \qquad\qquad V_R \to \frac{1}{2}c_{\varphi ud}v^2/\Lambda^2$$
$$g_L \to \sqrt{2}c_{uW}v^2/\Lambda^2 \qquad\qquad g_R \to \sqrt{2}c_{dW}v^2/\Lambda^2 \qquad (3.12)$$

Although anomalous couplings capture most of the same physics, the advantages of
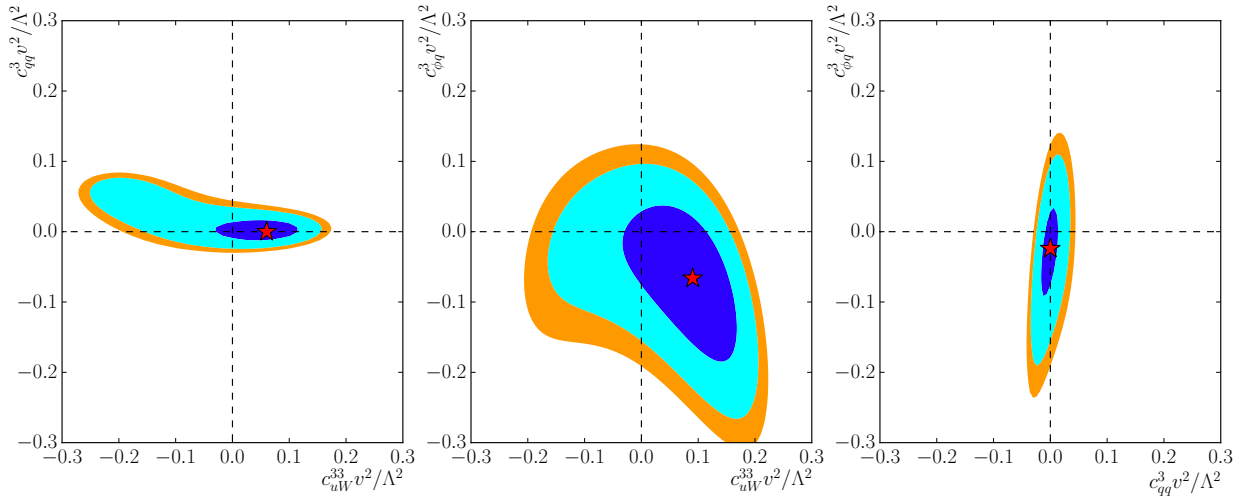
Figure 3.9: *Marginalised 68%, 95% and 95% confidence intervals on dimension-six operators in single top production.*

using higher-dimensional operators are manifold. Firstly, the power-counting arguments of the previous paragraph that allowed us to reject the operators $\mathcal{O}_{dW}$, $\mathcal{O}_{\varphi ud}$ at order Λ^{-2} would not be clear in an anomalous coupling framework. In addition, the four-quark operator $\mathcal{O}_{qq}^{(3)}$ in Eq. (3.9) can have a substantial effect on single-top production, but this can only be captured by an EFT approach. For a detailed comparison of these approaches, see e.g. Ref. [329]. The 95% confidence limits on these operators from single top production are shown in Fig. (3.10), along with those operators previously discussed in top pair production.

Let us compare these results to our findings of Section 3.4.1. The bounds on operators from top pair production are typically stronger. The so-called chromomagnetic moment operator $\mathcal{O}_{uG}$ is also tightly constrained, owing to its appearance in both the $q\bar{q}$ and gg channels, i.e. it is sensitive to both Tevatron and LHC measurements. For the four-quark operators, the stronger bounds are typically on the c_i^1-type. This originates from the more pronounced effect on kinematic distributions that they have. The phenomenology of the c_i^2-type operators is SM-like, and their effect becomes only visible in the tails of distributions.

The much wider marginalised bounds on these two operators come from the relative sign between their interference term and those of the other operators, which results in cancellations in the total cross-section that significantly widen the allowed ranges of c_i. With the exception of c_t, which strongly modifies the single top production cross-section, the individual bounds on the operator coefficients from single top production are typically weaker. This originates from the larger experimental uncertainties on single top production, that stem from the multitude of different backgrounds that contaminate this process,
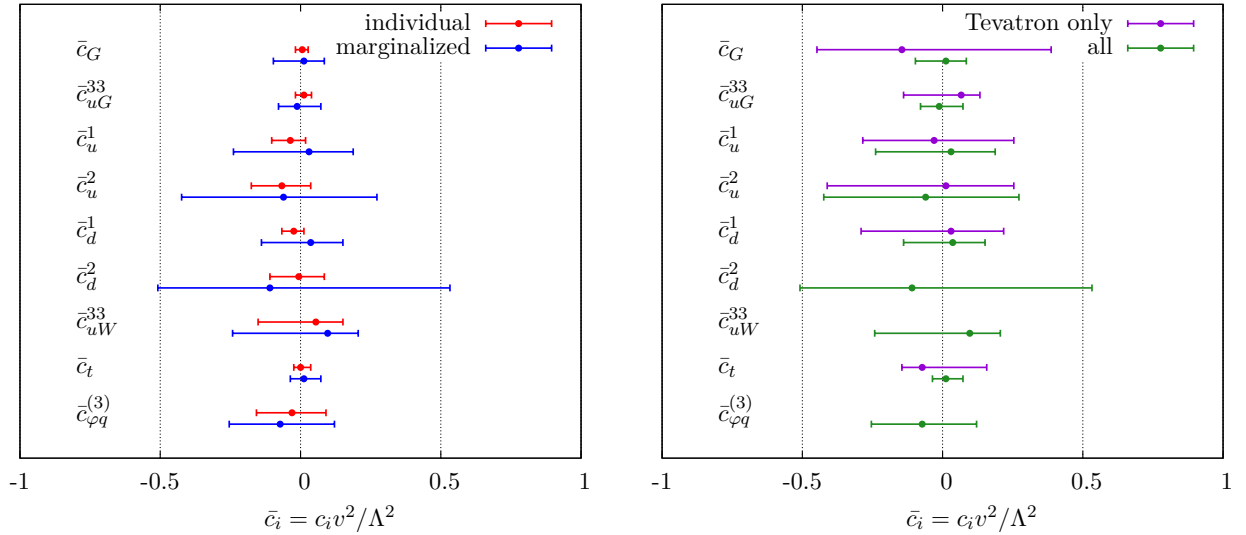
Figure 3.10: *Left: Individual (red) and marginalised (blue) 95% confidence intervals on dimension-six operators from top pair production and single top production (bottom three). Right: Marginalised 95 % bounds considering all data from LHC and Tevatron (green) vs Tevatron only (purple).*

particularly top pair production. For the Tevatron datasets this is particularly telling: the few measurements that have been made, with no differential distributions, combined with the large error bars on the available data, mean that two of the three operators are not constrained at dimension-six[§]. Still, as before, excellent agreement with the SM is observed.

In addition to single-top production, the operator $\mathcal{O}_{uW}$ may be constrained by distributions relating to the kinematics of the top quark decay. The matrix element for hadronic top quark decay $t \rightarrow Wb \rightarrow bqq'$, for instance, is equivalent to that for t-channel single top production via crossing symmetry, so decay observables provide complementary information on this operator. We will discuss the bounds obtainable from decay observables in Section 3.4.4.

3.4.3 Associated production

In addition to top pair and single top production, first measurements have been reported [304–306] of top pair production in association with a photon and with a Z boson ($t\bar{t}\gamma$ and $t\bar{t}Z$)[¶]. The cross-section for these processes are considerably smaller, and statis-

---

[§]Our bounds on these two operators are of the same order, but wider, than a pre-LHC phenomenological study [257], owing to larger experimental errors than estimated there.

[¶]Early measurements of top pair production in association with a W has also been reported by ATLAS and CMS, but the experimental errors are too large to say anything meaningful about new physics therein; the measured cross-sections are still consistent with zero.
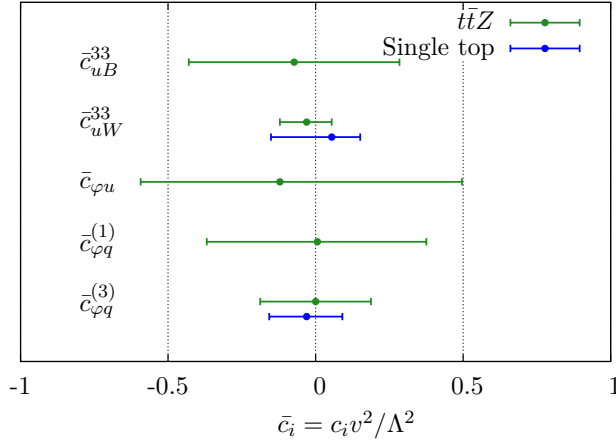
Figure 3.11: *Individual 95% confidence intervals for the operators of Eq. (3.13) from $t\bar{t}\gamma$ and $t\bar{t}Z$ production (green) and in the two cases where there is overlap, from single top measurements (blue).*

tical uncertainties currently dominate the quoted measurements. Still, they are of interest because they are sensitive to a new set of operators not previously accessible, corresponding to enhanced top-gauge couplings which are ubiquitous in simple W' and Z models, and which allow contact to be made with electroweak observables. The operator set for $t\bar{t}Z$, for instance, contains the 6 top pair operators in Eq. (3.8), plus the following

$$
\begin{aligned}
\mathcal{L}_{\text{D6}} \supset \; & \frac{c_{uW}}{\Lambda^2}(\bar{q}\sigma^{\mu\nu}\tau^I u)\,\tilde{\varphi}\,W^I_{\mu\nu} + \frac{c_{uB}}{\Lambda^2}(\bar{Q}\sigma^{\mu\nu}u)\,\tilde{\varphi}\,B_{\mu\nu} + \frac{c^{(3)}_{\varphi q}}{\Lambda^2}i(\varphi^\dagger \overleftrightarrow{D}^I_\mu \varphi)(\bar{Q}\gamma^\mu\tau^I Q) \\
& + \frac{c^{(1)}_{\varphi q}}{\Lambda^2}i(\varphi^\dagger \overleftrightarrow{D}_\mu \varphi)(\bar{Q}\gamma^\mu Q) + \frac{c_{\varphi u}}{\Lambda^2}(\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi)(\bar{u}\gamma^\mu u)\,.
\end{aligned}
\tag{3.13}
$$

There is therefore overlap between the operators contributing to associated production, and those contributing to both top pair and single top. In principle, one should include all observables in a global fit, fitting all coefficients simultaneously. However, the low number of individual $t\bar{t}V$ measurements, coupled with their relatively large uncertainties, means that they do not have much effect on such a fit. Instead, we choose to present individual constraints on the operators from associated production alone, comparing these with top pair and single top in what follows. For the former, we find that the constraints on the operators of Eq. (3.13) obtained from $t\bar{t}\gamma$ and $t\bar{t}Z$ measurements are much weaker than those obtained from top pair production, therefore we do not show them here. The constraints on the new operators of Eq. (3.13) are displayed in Fig. 3.11. It is interesting to note that the constraints from associated production measurements are comparable with those from single top production, despite the relative paucity of the former.
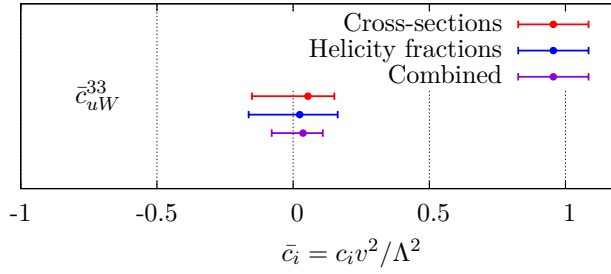
Figure 3.12: *95% bounds on the operator $\mathcal{O}_{uW}$ obtained from data on top quark helicity fractions (blue) vs. single top production cross-sections (red), and both sets of measurements combined (purple).*

3.4.4 Decay observables

This completes the list of independent dimension-six operators that affect top quark production cross-sections. However, dimension-six operators may also contribute (at interference level) to observables relating to top quark decay. Top quarks decay almost 100% of the time to a W and b quark. The fraction of these events which decay to W-bosons with a given helicity: left-handed, right-handed or zero-helicity, can be expressed in terms of helicity fractions, which for leading order with a finite b-quark mass were shown in Eq. (1.87):

$$
\begin{aligned}
F_0 &= \frac{(1-y^2)^2 - x^2(1+y^2)}{(1-y^2)^2 + x^2(1-2x^2+y^2)} \\
F_L &= \frac{x^2(1-x^2+y^2) + \sqrt{\lambda}}{(1-y^2)^2 + x^2(1-2x^2+y^2)} \\
F_R &= \frac{x^2(1-x^2+y^2) - \sqrt{\lambda}}{(1-y^2)^2 + x^2(1-2x^2+y^2)}
\end{aligned}
\tag{3.14}
$$

where $x = M_W/m_t$, $y = m_b/m_t$ and $\lambda = 1 + x^4 + y^4 - 2x^2y^2 - 2x^2 - 2y^2$. As noted in Ref. [217], measurements of these fractions can be translated into bounds on the operator $\mathcal{O}_{uW}$. The operator $\mathcal{O}_{\varphi q}^{(3)}$ cannot be accessed in this way, since its only effect is to rescale the Wtb vertex $V_{tb}^2 \rightarrow V_{tb}\left(V_{tb} + v^2 c_{\varphi q}^{(3)}/\Lambda^2\right)$, therefore it has no effect on event kinematics. The desirable feature of these quantities is that they are relatively stable against higher order corrections, so the associated scale uncertainties are small. The Standard Model NNLO estimates for these are: $\{F_0, F_L, F_R\} = \{0.687 \pm 0.005, 0.311 \pm 0.005, 0.0017 \pm 0.0001\}$ [146], i.e. the uncertainties are at the per mille level. It is interesting to ask whether the bound obtained on $\mathcal{O}_{uW}$ in this way is stronger than that obtained from cross-section measurements. In Fig. 3.12 we show the constraints obtained in each way. Although they are in excellent agreement with each other, cross-section information gives a slightly stronger bound, mainly due to the larger amount of data available, but also due to the
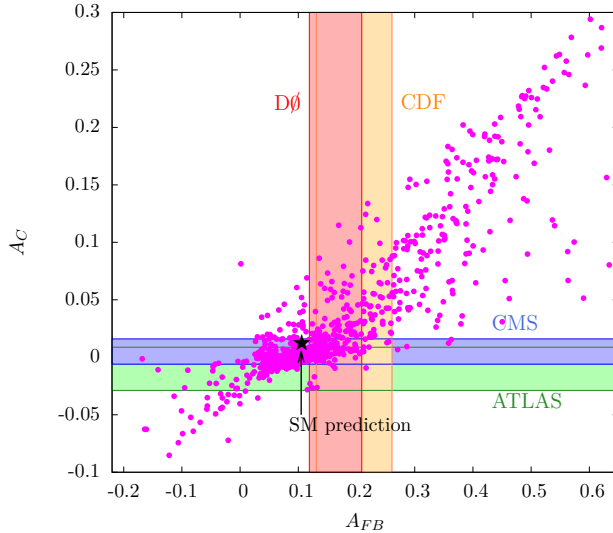
Figure 3.13: *Results of a 1000 point parameter space scan over -10 $TeV^{-2} < c_{u,d}^{1,2}/\Lambda^2 < 10\,TeV^{-2}$ overlaid with the most up to date measurements of A_{FB} and A_{C}, showing clearly the correlation between them.*

large experimental uncertainties on F_i. Still, these measurements provide complementary information on the operator $\mathcal{O}_{uW}$, and combining both results in a stronger constraint than either alone, as expected.

3.4.5 Charge asymmetries

Asymmetries in the production of top quark pairs have received a lot of attention in recent years, particularly due to an apparent discrepancy between the Standard Model prediction for the so-called 'forward-backward' asymmetry A_{FB} in top pair production of Eq. (1.77)

$$A_{\mathrm{FB}} = \frac{N(\Delta y > 0) - N(\Delta y < 0)}{N(\Delta y > 0) + N(\Delta y < 0)} \tag{3.15}$$

where $\Delta y = y_t - y_{\bar{t}}$, and a measurement by CDF [330]. This discrepancy was most pronounced in the high invariant mass region, pointing to potential TeV-scale physics at play. However, recent work has cast doubts on its significance for two reasons: Firstly, an updated analysis with higher statistics [296] has slightly lowered the excess. Secondly, a full NNLO QCD calculation [116] of A_{FB} showed that, along with NLO QCD + electroweak calculations [111,115,331] the radiative corrections to A_{FB} are large. The current measurements for the inclusive asymmetry are now consistent with the Standard Model within 2σ. Moreover, the DØ experiment reports [297] a high-invariant mass measurement *lower* than the SM prediction. From a new physics perspective, it is difficult to accommodate all of this information in a simple, uncontrived model without tension.

Still, in an effective field theory approach, deviations from the Standard Model prediction of A_{FB} take a very simple form. A non-zero asymmetry arises from the difference of four-quark operators:

$$A_{\mathrm{FB}} = (c_u^1 - c_u^2 + c_d^1 - c_d^2)\frac{3s\beta}{4g_s^2\Lambda^2(3 - \beta^2)}, \qquad (3.16)$$

where $\beta = \sqrt{1 - s/4m_t^2}$ is the velocity of the $t\bar{t}$ system[‖]. Combining this inclusive measurement with differential measurements such as $dA_{\mathrm{FB}}/dm_{t\bar{t}}$ allows simultaneous bounds to be extracted on all four of these operators. Therefore it is instructive to compare the bounds obtained on $c_{u,d}^{1,2}$ from charge asymmetries to those obtained from $t\bar{t}$ cross-sections. Again it is possible to (indirectly) investigate the complementarity between Tevatron and LHC constraints. Though the charge symmetric initial state of the LHC does not define a 'forward-backward' direction, a related charge asymmetry can be (Eq. 1.78) defined as:

$$A_C = \frac{N(\Delta|y| > 0) - N(\Delta|y| < 0)}{N(\Delta|y| > 0) + N(\Delta|y| < 0)}, \qquad (3.17)$$

making use of the fact that tops tend to be produced at larger rapidities than antitops. This asymmetry is diluted with respect to A_{FB}, however. The most up-to-date SM prediction is $A_C = 0.0123 \pm 0.005$ [111] for $\sqrt{s} = 7$ TeV. The experimental status of these measurements is illustrated in Fig. 3.13. The inclusive measurements of A_{FB} are consistent with the SM expectation, as are those of A_C. The latter, owing to large statistical errors, are also consistent with zero, however, so this result is not particularly conclusive. Since these are different measurements, it is also possible to modify one without significantly impacting the other. Clearly they are correlated, however, as evidenced in Fig. 3.13, where the most up to date measurements of A_{FB} and A_C are shown along with the results of a 1000 point parameter space scan over the four-quark operators. This highlights the correlation between the two observables: non-resonant new physics which causes a large A_{FB} will also cause a large A_C, provided it generates a dimension-six operator at low energies.

We have used both inclusive measurements of the charge asymmetries A_C and A_{FB}, and measurements as a function of the top pair invariant mass $m_{t\bar{t}}$ and rapidity difference $|y_{t\bar{t}}|$. In addition, ATLAS has published measurements of A_C with a longitudinal 'boost' of the $t\bar{t}$ system: $\beta = (|p_t^z + p_{\bar{t}}^z|)/(E_t + E_{\bar{t}}) > 0.6$, which may enhance sensitivity to new physics contributions to A_C, depending on the model [334]. Since $A_{\mathrm{FB}} = 0$ at leading-order in the SM, it is not possible to define a K-factor in the usual multiplicative sense. Instead we take higher-order QCD effects into account by adding the NNLO QCD prediction to the dimension-six terms. In the case of A_C, we normalise the small (but non-zero) LO

[‖]Contributions to A_{FB} also arise from the normalisation of A_{FB} and the dimension-six squared term [113, 332, 333], which we keep, as discussed in Sections 3.3 and 4.
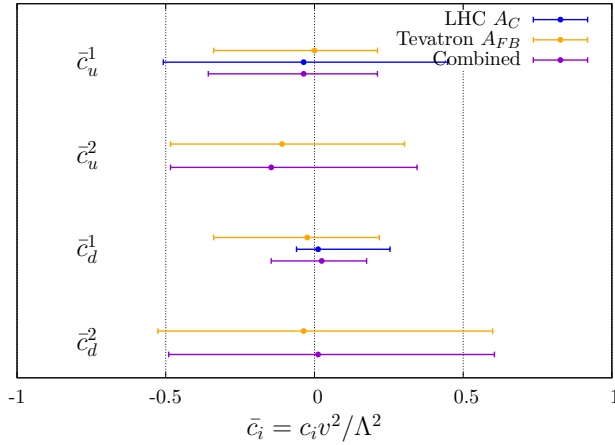
Figure 3.14: *Marginalised 95% confidence intervals on top pair four quark operators from charge asymmetries at the LHC and Tevatron.*

QCD piece, to the NLO prediction, which has been calculated with a Monte Carlo and cross-checked with a dedicated NLO calculation [111].

The above asymmetries have been included in the global fit results presented in Fig. 3.18. However, it is also interesting to see what constraints are obtained on the operators from asymmetry data alone. To this end, the 95% confidence intervals on the coefficients of the operators $\mathcal{O}_{u,d}^{1,2}$ from purely charge asymmetry data are shown in Fig. 3.14. Unsurprisingly, the bounds are much weaker than for cross-section measurements in Fig. 3.10, with the $\mathcal{O}_i^2$-type operators unconstrained by LHC data alone. Despite the small discrepancy between the measured A_{FB} and its SM value, this does not translate into a non-zero Wilson coefficient; as before, all operators are zero within the 95% confidence intervals. At 13 TeV, the asymmetry A_{C} will be diluted even further, due to the increased dominance of the $gg \rightarrow t\bar{t}$ channel, for which $A_C = 0$. It is therefore possible that charge asymmetry measurements (unlike cross-sections) will not further tighten the bounds on these operators during LHC Run II.

3.4.6   Contribution of individual datasets

Counting each bin independently, there are a total of 234 measurements entering our fit, giving a total χ^2 of 206.1 at the best fit point. It is instructive to examine how this is distributed across the individual datasets. We quantify this by calculating the χ^2 per bin between the data and the global best fit point, as shown in Fig. 3.15. Overall, excellent agreement is seen across the board, with no measurement in obvious tension with any other. The largest single contributors to the χ^2 come from the rapidity distributions in top pair production. It has been known for some time that there is some tension between data and Monte Carlo generators for this observable, especially in the forward region (see e.g.
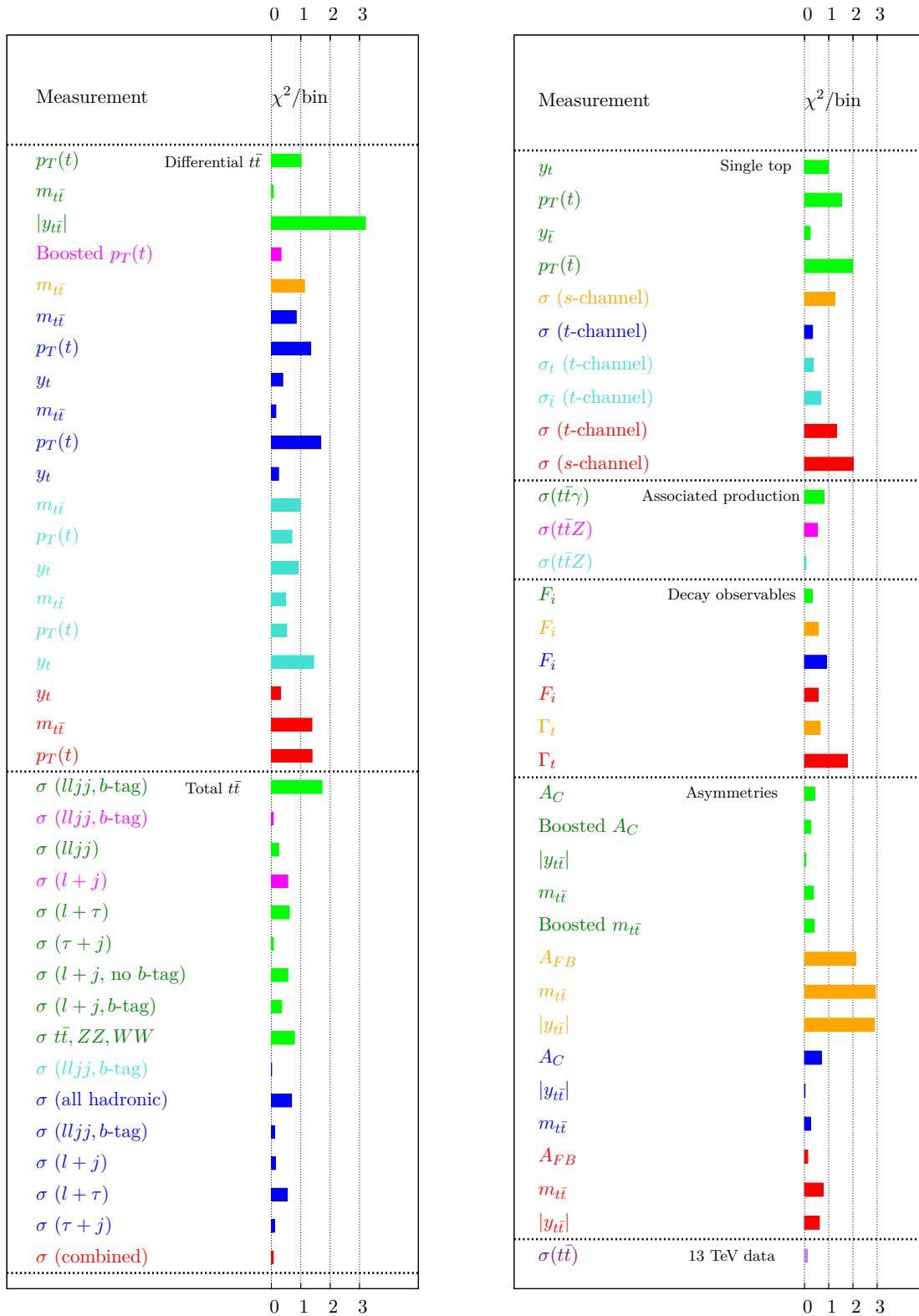
Figure 3.15: χ^2 *per bin between measurement and the interpolated best fit point, for measurements considered in this fit. Colours: Green: ATLAS 7 TeV, Magenta: ATLAS 8 TeV, Blue: CMS 7 TeV, Turquoise: CMS 8 TeV, Red: DØ, Orange: CDF, Purple: CMS 13 TeV.*

the rapidity distributions in Ref. [292]). It is quite likely that this discrepancy stems from the QCD modelling of the event kinematics, rather than potential new physics. Moreover, in a fit with this many measurements, discrepancies of this magnitude are to be expected on purely statistical grounds.

At the level of total $t\bar{t}$ cross-sections, the vanishingly small contributions to the χ^2 stem from the simple fact that the total rate is well-described by the SM. Single top production measurements are also in good agreement with the SM. The associated production processes $tt\gamma$ and ttZ, along with the charge asymmetry measurements from the LHC, have a very small impact on the fit, owing to the large statistical uncertainties on the current measurements. For the former, this situation will improve in Run II, for the latter the problem will be worse. The forward-backward asymmetry measurements from CDF remain the most discrepant dataset used in the fit, even though the inclusive asymmetry is in good agreement with the NNLO SM asymmetry, because the NNLO corrections shrink the scale uncertainty band whilst not enlarging the central value, thus enhancing the CDF excess (see Ref. [116]).

3.5 Validity of the EFT approach

As we have just shown, collider measurements can be used to extract bounds on Wilson coefficients in a completely model-independent way. However, if the corresponding dimension-six operators are to be interpreted as the leading terms in a consistent effective field theory, then care must be taken to ensure the constraints are valid. The first question one might ask is if the dimension-six truncation is valid, and more broadly one can ask if the measurements used in the fit are probing kinematic regimes that make the entire effective description invalid.

3.5.1 Impact of quadratic terms

As discussed in section 2.6, at the level of observables adding $D = 6$ operators to the Standard Model Lagrangian amounts to replacing the Standard Model matrix element with

$$|\mathcal{M}_{\text{full}}|^2 = |\mathcal{M}_{\text{SM}}|^2 + 2\Re\mathcal{M}_{\text{D6}}^*\mathcal{M}_{\text{SM}} + |\mathcal{M}_{\text{D6}}|^2. \tag{3.18}$$

The last term on the right-hand side is of order $\mathcal{O}(1/\Lambda^4)$. However, to have a fully consistent description of an observable at this order, one should include *all* terms proportional to $\mathcal{O}(1/\Lambda^4)$, i.e. also include the interference terms between the Standard Model and dimension-eight operators. In the strictest interpretation of the effective Lagrangian
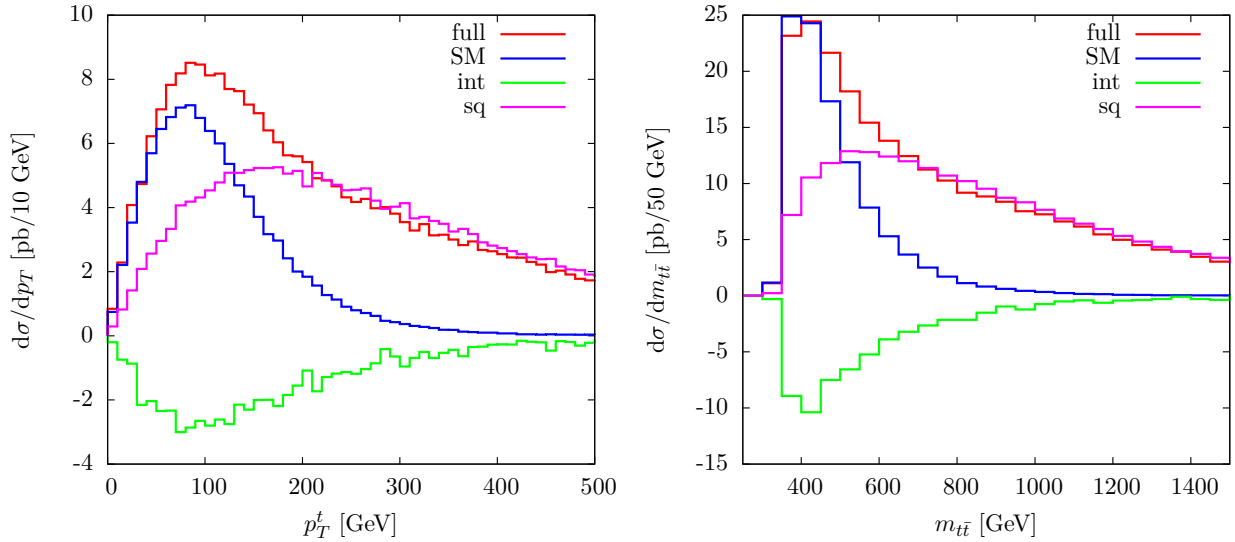
106

Figure 3.16: *Distributions in top quark pair production for a given point in the parameter space of $t\bar{t}$ Wilson coefficients considered in this fit, namely $\{c_i\} = \{c_G, c_{uG}^{33}, c_u^1, c_u^2, c_d^1, c_d^2\} = \{0.01, 0.01, 0, 0.9, 0, -9.2, -9.2\}$ TeV^{-2}. In blue is the SM prediction, in green the pure dimension-six contribution, in pink the pure squared term, and in red the sum of the three. Subtracting off the square terms leads to negative cross-sections in the intermediate mass range. All distributions are leading-order in α_s.*

approach, neglecting these terms renders the EFT description of such an observable meaningless. This is an overly restrictive viewpoint, however, as there exist several prescriptions for ensuring that the dimension-six approximation is under control.

One might, for example, restrict the parameter space of Wilson coefficients to regions for which the linear interference terms dominate over the quadratic terms. This ensures that all constraints are valid in a dimension-six effective field theory interpretation. However, it may reduce the sensitivity to the operators, and weaken (or wipe out altogether) the obtained constraints. Moreover, there is an ambiguity in how tolerant of the squared terms one should be: should one, for example, only cut off the parameter space when they become *larger* than the interference terms, in which case their effects will still be considerable, or when they reach 10, 20 or 50% of the linear piece? There is no first principles answer to this.

Alternatively, one could subtract off the squared terms altogether, either at the level of matrix elements or of observables. By definition this ensures that the constraints can be interpreted in terms of a dimension-six EFT. This treatment, however, lacks a clear physical motivation. If a non-zero Wilson coefficient were explicitly measured, it would have to correspond to the low energy limit of some ultraviolet completion. Matching the EFT constraints onto a specific UV model can be done in a general way, and will determine

if the data favours a 'natural' region of that model's parameter space. However, subtracting off quadratic effects will in general disrupt the accuracy of the matching procedure, as the most dominant effects of a particular model on a certain observable may be from the non-interfering piece[**]. Furthermore, subtracting off some contributions to the cross-section might lead to unphysical effects in distributions, as shown in Fig. 3.16.

By looking at the range $p_T^t \gtrsim 200$ GeV, and $m_{t\bar{t}} \gtrsim 600$ GeV, we see that considering only the interference term leads to negative predictions for the differential cross-section: a clearly unacceptable result. Here the terms proportional to $\mathcal{O}(1/\Lambda^4)$ keep the cross-section physically meaningful. One could take this as evidence that the dimension-six approximation is then not valid at all, because the squared term is dominating. However, if one would like to compare the EFT constraints to those obtained in a specific new physics model, it is necessary to include these to ensure an accurate matching condition. The importance of keeping the squared term or not is then ultimately a model-dependent question.

Finally, one could simply adopt the pragmatic approach of viewing the dimension-six framework not as the leading part of a consistent effective theory, but as a model-independent way of parameterising how well the Standard Model describes the data. If a non-zero Wilson coefficient was to be measured, it would still be evidence for new physics, even if it could not be simply linked to a particular new physics model. This is the approach we adopt in this fit, and we leave interpretational issues aside.

Even if one considers a fit with quadratic terms completely removed from all observables, $\mathcal{O}(\Lambda^{-4})$ terms manifest in another way when likelihood contours are drawn. Consider the likelihood function of Eq. (3.6), considering just one bin and one operator for simplicity. For a linear fitting function f, it will have a polynomial dependence on c_i proportional to

$$\chi^2(c_i) = \frac{(f(c_i) - E)^2}{\Delta^2} \sim c_i + c_i^2 - E. \tag{3.19}$$

To ensure the χ^2 has a local minimum, the squared term must be kept. This provides another argument in favour of keeping the quadratic terms throughout the fit.

3.5.2 Overflow bins

Related, but not identical, to the question of whether one should omit or keep the quadratic terms, is whether one should worry about events in the tails of distributions that might invalidate the EFT treatment. For inclusive observables this is less of a problem, because they tend to be dominated by electroweak scale thresholds, well within the valid region of the phase space. However, by naïve power counting, the convergence of the EFT expansion

[**]See also Refs. [335, 336] for a discussion of this.

rests on the two conditions

$$\frac{g_*^2 v^2}{\Lambda^2} < 1 \quad \text{and} \quad \frac{E^2}{\Lambda^2} < 1, \tag{3.20}$$

where g_* is a generic new physics coupling, v is the Higgs vev, and E is the maximum energy scale probed by the process. The first condition can be ensured for any weakly coupled UV completion. The second condition is troublesome when differential distributions are included in the fit, however. The final bin in experimentally measured distributions, such as those in Fig. 3.6, typically contains not only events in that phase space region, but also so-called overflow events to the right of the plot. If the experiment has not included information on the maximum momentum transfer probed in the published dataset, i.e. the maximum value of the 'overflow' entry to the right of the plot, it is difficult to consistently interpret the resulting bounds on the Wilson coefficients in an underlying UV model, because those overflow events may violate the power counting conditions of Eq. (3.20). As a test of how much pull they have on the fit, in Fig. 3.17 we plot the 1-dimensional likelihood distributions (equivalent to $\Delta\chi^2 = \chi^2 - \chi^2_{min}$) for the Wilson coefficients relevant for top pair production, considering the full $t\bar{t}$ dataset, and omitting the overflow bins in the kinematic distributions.

The differences in the constraints are small, typically at the order of a few percent, showing that the fit is not unduly biased by phase space points that undermine the validity of the effective field theory. One may take the approach of omitting these overflow bins altogether, to ensure control over the scales involved in the fitted measurements. The subsequent limits are then slightly weaker, due to reduced statistical power and sensitivity to the operators. However, the exclusion of certain data points undermines the 'global' nature of a global fit, so they are included for full generality. Indeed, this is again a model-dependent question. To illustrate this, in the next section we will convert our EFT constraints onto specific UV models.

Before doing this, a final comment is in order about Fig. 3.17. The limit setting shown there, is obtained from a likelihood ratio test, rather than the raw χ^2 we employ elsewhere in the fit. This is so that the 2σ constraints for both datasets can be easily shown on the same plot, as they both correspond to $\Delta\chi^2(c_i) < 4$. For the χ^2 test both datasets have different numbers of degrees of freedom, corresponding to the number of input measurements they contain. However, the same results apply in this case as well. The limits obtained in the latter approach (i.e. the one we adopt in the rest of the fit) are actually weaker. Without a compelling reason to adopt either approach, it thus seems sensible to take the more conservative option.
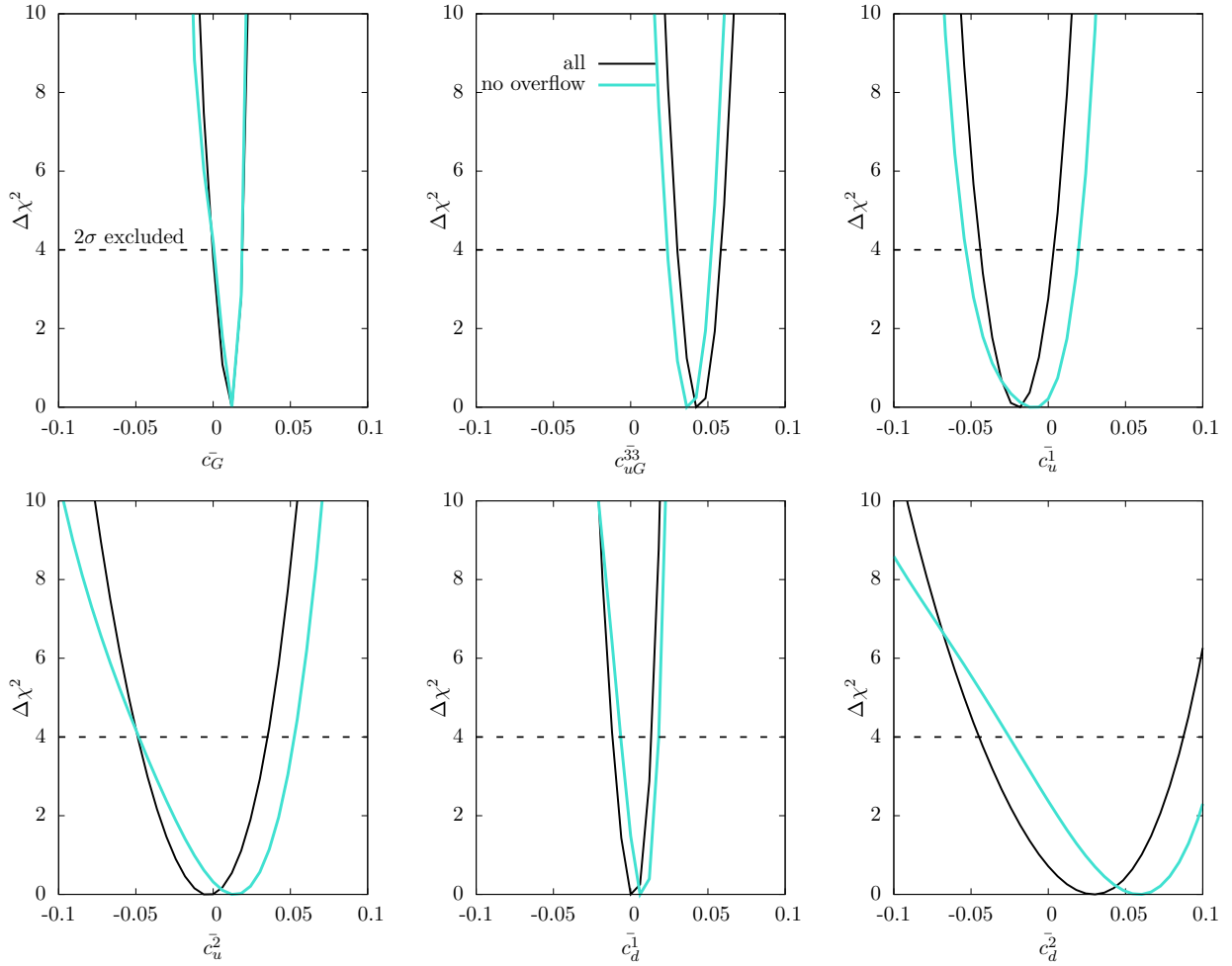
Figure 3.17: *Individual one-dimensional likelihood plots for the top pair Wilson coefficients considered in this fit, containing all differential top pair measurements (black), and omitting the final bin in the $m_{t\bar{t}}$ and p_T^t distributions (turquoise).*

3.5.3 Constraining UV models

As an illustration of the wide-ranging applicability of EFT techniques, we conclude by matching our effective operator constraints to the low-energy regime of some specific UV models. These models do not necessarily represent concrete UV scenarios, but serve as illustrative examples of how EFT constraints could map onto the parameter space of more fundamental theories.

s-channel axigluon:

Considering top pair production, one can imagine the four operators of Eq. (2.62) as being generated by integrating out a heavy s-channel resonance which interferes with the QCD $q\bar{q} \to t\bar{t}$ amplitude. One particle that could generate such an interference is the so-called

axigluon. These originate from models with an extended strong sector with gauge group $SU(3)_{c1} \times SU(3)_{c2}$ which is spontaneously broken to the diagonal subgroup $SU(3)_c$ of QCD. In the most minimal scenario, this breaking can be described by a non-linear sigma model

$$\mathcal{L} = -\frac{1}{4}G_{1\mu\nu}G_1^{\mu\nu} - \frac{1}{4}G_{2\mu\nu}G_2^{\mu\nu} + \frac{f^2}{4}\mathrm{Tr}D_\mu\Sigma D^\mu\Sigma^\dagger \quad , \quad \Sigma = \exp\left(\frac{2i\pi^a t^a}{f}\right) \quad , \quad a = 1, ..., 8. \tag{3.21}$$

Here π^a represent the Goldstone bosons which form the longitudinal degrees of freedom of the colorons, giving them mass, t^a are the Gell-Mann matrices, and f is the symmetry breaking scale. The nonlinear sigma fields transform in the bifundamental representation of $SU(3)_{c1} \times SU(3)_{c2}$:

$$\Sigma \to U_L\Sigma U_R^\dagger \quad , \quad U_L = \exp\left(\frac{i\pi^a\alpha_L^a}{f}\right) \quad , \quad U_R = \exp\left(\frac{i\pi^a\alpha_R^a}{f}\right). \tag{3.22}$$

The physical fields are obtained by rotating the gauge fields G_1 and G_2 to the mass eigenstate basis

$$\begin{pmatrix} G_{1\mu}^a \\ G_{2\mu}^a \end{pmatrix} = \begin{pmatrix} \cos\theta_c & -\sin\theta_c \\ \sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} G_\mu^a \\ C_\mu^a \end{pmatrix}, \tag{3.23}$$

where the mixing angle θ_c is defined by

$$\sin\theta_c = \frac{g_{s1}}{\sqrt{g_{s1}^2 + g_{s2}^2}}. \tag{3.24}$$

The case of an axigluon corresponds to maximal mixing $\theta = \pi/4$, i.e. $g_{s1}^2 = g_{s2}^2 = g_s^2/2$. Taking the leading-order interference with the SM amplitude for $q\bar{q} \to t\bar{t}$, in the limit $s \ll M_A^2$, we find that the axigluon induces the dimension-six operators

$$\frac{c_u^1}{\Lambda^2} = \frac{g_s^2}{M_A^2}, \qquad \frac{c_d^1}{\Lambda^2} = \frac{5g_s^2}{4M_A^2}, \qquad \frac{c_u^2}{\Lambda^2} = \frac{c_d^2}{\Lambda^2} = \frac{2g_s^2}{M_A^2}. \tag{3.25}$$

Substituting the marginalised constraints on the 4-quark operators, we find this translates into a lower bound on an axigluon mass. $M_A \gtrsim 1.4$ TeV at the 95% confidence level. Since this mass range coincides with the overflow bin of figure 3.6, this bound creates some tension with the validity of the EFT approach in the presence of resonances in the $t\bar{t}$ spectrum (for a general discussion see Ref. [320, 337, 338]); at this stage in the LHC programme indirect searches are not sensitive enough to compete with dedicated searches.

s-channel W':

Turning our attention to single top production, we consider the example of the operator $\mathcal{O}_{qq}^{(3)}$ being generated by a heavy charged vector resonance (W') which interferes with the

SM amplitude for s-channel single top production: $u\bar{d} \to W \to t\bar{b}$. The most general Lagrangian for such a particle (allowing for left and right chiral couplings) is (see e.g. Ref. [339]):

$$\mathcal{L} = \frac{1}{2\sqrt{2}} V_{ij} g_{W'} \bar{q}_i \gamma_\mu (f_{ij}^R (1 + \gamma^5) + f_{ij}^L (1 - \gamma^5)) W^\mu q_j + h.c. \tag{3.26}$$

We take the generic coupling $g_{W'} = g_{SM}$. Since we are considering the interference term only, which must have the same $(V - A)$ structure as the SM, we can set $f^R = 0$. Considering the tree-level interference term for between the diagrams for $u\bar{d} \to W', W' \to t\bar{b}$, and taking the limit $s \ll M_W'^2$ (we also work in the narrow-width approximation $\Gamma_W, \Gamma_{W'} \ll M_W, M_{W'}$), we find

$$\frac{c_{qq}^{3,1133}}{\Lambda^2} = \frac{g^2}{4M_{W'}^2}, \tag{3.27}$$

which, using our global constraint on $\mathcal{O}_t$, translates into a bound $M_{W'} \gtrsim 1.2$ TeV.

These bounds are consistent with, but much weaker than, constraints from direct searches for dijet resonances from ATLAS [340, 341] and CMS [342], which report lower bounds of $\{M_A, M_{W'}\} > \{2.72, 3.32\}$ TeV and $\{M_A, M_{W'}\} > \{2.2, 3.6\}$ TeV respectively. It is unsurprising that these dedicated analyses obtain stronger limits, given the generality of this fit. Again this energy range is resolved by the fit thus in principle invalidating the EFT approach to obtain Eq. (3.27). Nonetheless, these bounds provide an interesting comparison of our numerical results, whilst emphasising that for model-specific examples, direct searches for high-mass resonances provide stronger limits than general global fits.

3.6 Summary

In this chapter, we performed a global fit of top quark effective field theory to experimental data, including all constrainable operators at dimension six. For the operators, we use the 'Warsaw basis' of Ref. [199], which has also been widely used in the context of Higgs and precision electroweak physics. We use data from the Tevatron and LHC experiments, including LHC Run II data, up to a centre of mass energy of 13 TeV. Furthermore, we include fully inclusive cross-section measurements, as well as kinematic distributions involving both the production and decay of the top quark. Counting each bin independently, the total number of observables entering the fit is 234, with a total of 12 contributing operators. Constraining the coefficients of these operators is then a formidable computational task. To this end we use the parametrisation methods in the PROFESSOR framework, first developed in the context of Monte Carlo generator tuning [307], and discussed here in Section 3.3.

We perform a χ^2 fit of theory to data, including appropriate correlation matrices where these have been provided by the experiments. We obtain bounds on the Wilson coeffi-
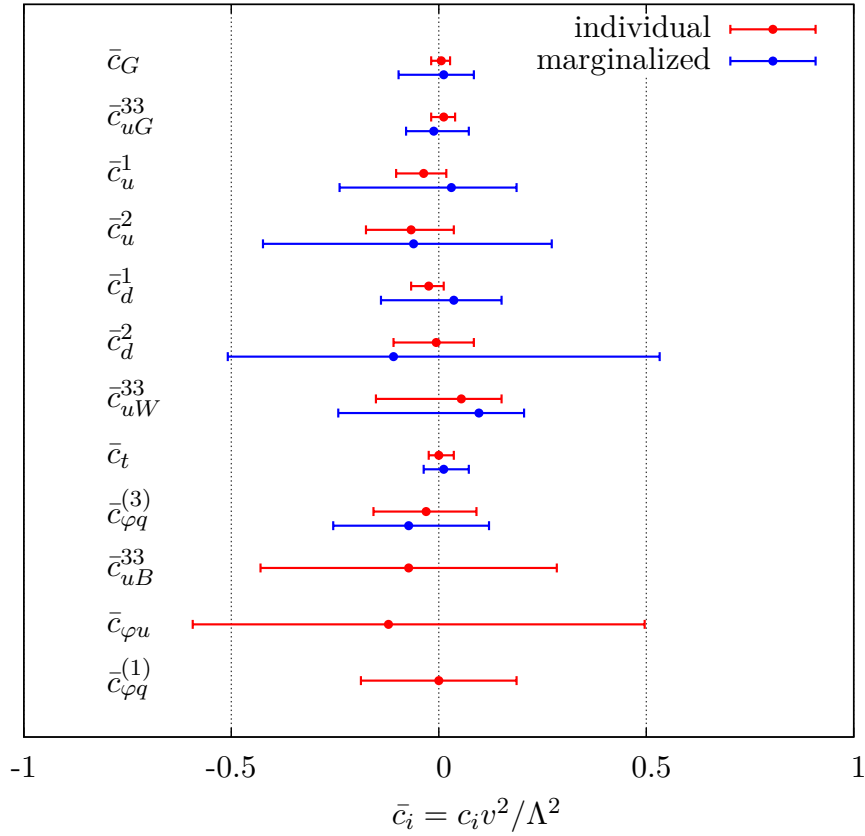
Figure 3.18: *95% confidence intervals for the dimension-six operators that we consider here, with all remaining operators set to zero (red) and marginalised over (blue). In cases where there are constraints on the same operator from different classes of measurement, the strongest limits are shown here. The lack of marginalised constraints for the final three operators is discussed in Section 3.4.3.*

cients of various operators contributing to top quark production and decay, summarised in Fig. 3.18, in two cases: (i) when all other coefficients are set to zero; (ii) when all other operators coefficients are marginalised over. The numerical values of these constraints are also shown in table 3.2.

The strongest constraints are on operators involving the gluon, as expected given the dominance of gluon fusion in top pair production at the LHC (for which there is more precise data). Four fermion operators are constrained well in general, with weaker constraints coming from processes whose experimental uncertainties remain statistically dominated (e.g. $t\bar{t}V$ production). We have quantified the interplay between the Tevatron and LHC datasets, as well as that between different measurement types (e.g. top pair, single top).

The results are all in agreement with the SM only hypothesis, with no tensions beyond the 95% confidence level, which is perhaps to be expected given the lack of reported deviations in previous studies. However, the fact that this agreement is obtained, in a

Coefficient	Individual constraint	Marginalised constraint
$c_G v^2/\Lambda^2$	(—0.018, 0.027)	(—0.097, 0.085)
$c_{uG}^{33} v^2/\Lambda^2$	(—0.018, 0.039)	(—0.079, 0.073)
$c_u^1 v^2/\Lambda^2$	(—0.103, 0.018)	(—0.236, 0.188)
$c_u^2 v^2/\Lambda^2$	(—0.175, 0.036)	(—0.424, 0.272)
$c_d^1 v^2/\Lambda^2$	(—0.067, 0.121)	(—0.139, 0.151)
$c_d^2 v^2/\Lambda^2$	(—0.109, 0.085)	(—0.508, 0.533)
$c_{uW}^{33} v^2/\Lambda^2$	(—0.151, 0.151)	(—0.242, 0.206)
$c_t v^2/\Lambda^2$	(—0.024, 0.036)	(—0.036, 0.073)
$c_{\varphi q}^{(3)} v^2/\Lambda^2$	(—0.157, 0.091)	(—0.254, 0.121)
$c_{uB}^{33} v^2/\Lambda^2$	(—0.430, 0.284)	(—, —)
$c_{\varphi u} v^2/\Lambda^2$	(—0.593, 0.496)	(—, —)
$c_{\varphi q}^{(1)} v^2/\Lambda^2$	(—0.188, 0.188)	(—, —)

Table 3.2: *Numerical values of the individual and marginalised 95% confidence intervals on the operators presented here.*

wide global fit, is itself testament to the consistency of different top quark measurements, with no obvious tension between overlapping datasets. New data from LHC Run II is continuously appearing, and can be implemented in our fit framework in a systematic way. Still, there are several potential refinements of our analysis that can be made in order to improve the numerical constraints presented here, that go beyond simply replacing 8 TeV measurements with 13 TeV ones. A discussion of these issues is the subject of the next chapter.

4 Future prospects for top quark EFT

4.1 Introduction

In chapter 3 we performed a comprehensive global fit of the $D = 6$ operators that can influence top quark observables at hadron colliders to all the published top measurements from the Tevatron and Run I of the LHC. These constituted (predominantly) top pair production in various decay channels, as well as single top, associated vector boson production and observables from top quark decay. Despite the impressive statistical sample that entered the fit, the subsequent bounds on the studied Wilson coefficients are rather weak, pointing to values of Λ of order $\mathcal{O}(\lesssim 1\text{TeV})$, depending on the assumed size of the UV couplings. While one can take care to ensure that all bounds are consistent within an EFT formulation, for example by cutting out the high-mass 'overflow' bins in the differential distributions used, for which there is no control over scales, this is still a disappointingly low scale compared to the design mass reach of the LHC.

The wide allowed ranges for these operators stems not from a lack of sensitivity to the operators, but from the large experimental systematics and theory uncertainty bands from varying the scales and PDFs, and the more general problem of searching for precision deviations at a hadron collider, rather than 'bump hunts'. Still, given that we are at a very early stage of the full LHC programme, it is well-motivated to ask what improvements can be made over its lifetime as theory descriptions are improved and experimental error bars are shrank. We saw that, for the case of $t\bar{t}$ production, vast improvement could be achieved by adding differential distributions as well as total rates. Typically, however, the measurements used in the fit were based on standard top reconstruction techniques, which while providing good coverage of the low p_T threshold region, suffer from large statistical and systematic uncertainties in the high p_T tails, precisely the region where we want to be most sensitive to the effects of the operators.

Moreover, the distributions used were typically unfolded to parton level; that is, the final-state objects were corrected for detector effects and the actual measured cross-section in a fiducial volume of the detector extrapolated to the full phase-space, without cuts. This substantially eases the workflow of our fit, since the data can be compared directly to parton-level predictions without the need for the full parton shower, hadronisation and detector simulation chain to be implemented at each point in the parameter space. However, the extrapolation, which makes use of comparing to Monte Carlo simulations, necessarily biases the unfolded distributions towards SM-like shapes. It also introduces additional correlations between neighbouring bins which can broaden the χ^{2*}.

Both of these problems can be attacked by employing 'boosted' top reconstruction

---

*Experimental resolved measurements are now provided at hadron level as well.

techniques. Rather than the standard 'resolved' reconstruction techniques used in the analyses of chapter 3, which require the decay products from the top to be relatively well-separated in the detector, these are optimised for events in which the top is produced at very large p_T so that its decay products are collimated, and can be captured in a single large radius *fat* jet, in contrast to the typical one-to-one parton-jet matching of a resolved event reconstruction. This has the potential to dramatically increase our sensitivity to the high-p_T region. In addition, boosted reconstruction necessitates a hadron level description, so the model-dependence of the constraints induced by the unfolding procedure can also be mitigated. It is instructive to quantify how much they can improve the limits from the Run I fit. This is the subject of the first part of this chapter.

In the second part of this chapter, we move away from hadron colliders and study the role that future lepton colliders can play in this endeavour, focusing on the two most mature proposed colliders: the International Linear Collider (ILC) and Compact Linear Collider (CLIC). While in general lepton collider measurements will be sensitive to a different set of operators, there is overlap with LHC measurements, so that the sensitivities can be directly compared.

This chapter is structured as follows: In section 4.2, we discuss the improvements on the $t\bar{t}$ constraints that can be made by employing boosted jet substructure techniques. We analyse the importance of improving experimental systematics as well as collecting larger statistics, and the gain that can be made when theory uncertainties are improved beyond their present values. We also study the implications of our constraints for the reach of the LHC for generic (perturbative) UV completions. In section 4.3 we discuss the potential for improving the bounds on the electroweak operators in the top quark sector of the SMEFT, which can be accessed at hadron colliders through the process $pp \rightarrow t\bar{t}Z$. In section 4.4 we switch to lepton colliders, and compare the bounds on the operators of section 4.3 with the bounds that can be achieved with the forecasted capabilities of the ILC and CLIC colliders, before summarising in section 4.6.

4.2 Improving the fit with boosted reconstruction

Top pair production is (at leading order in α_s) a $2 \rightarrow 2$ process, so the relevant observables which span the partonic phase space are the momentum transfer $\hat{t}$ and the partonic centre-of-mass energy $\hat{s}$. All other observables are functions of these parameters, of which the top quark transverse momentum is the most crucial in determining the quality and efficiency of the boosted tagging approach which we will employ here [343–349]. As discussed in chapters 2 and 3, at leading order in the Standard Model EFT, the operators that contribute to top pair production are: the three-gluon vertex operator $\mathcal{O}_G$, the top chromomagnetic dipole moment operator $\mathcal{O}_{uG}$, as well as six four-quark operators $\mathcal{O}_{4q}$, which contribute
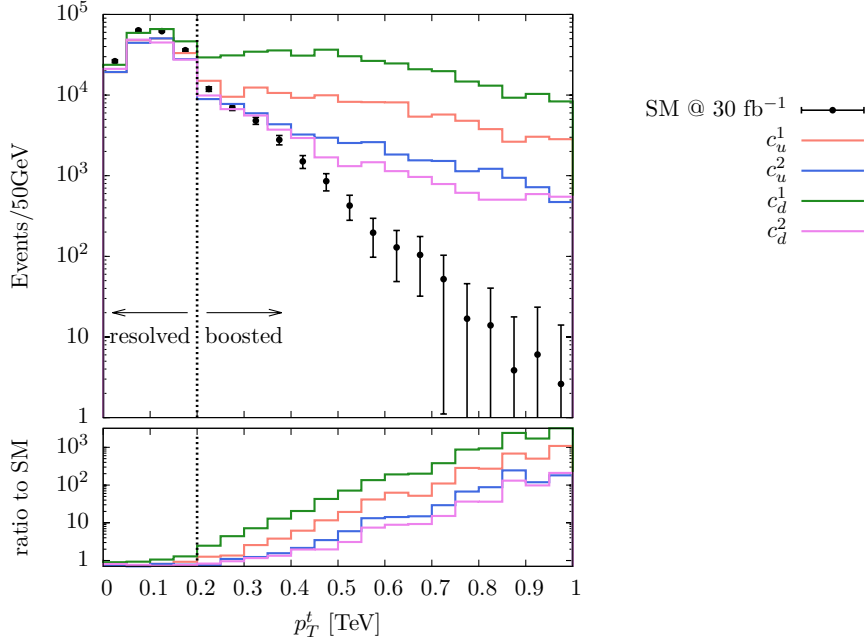
Figure 4.1: *Transverse momentum distributions for the reconstructed hadronic top quark candidate. The bars represent 30 fb^{-1} of pseudodata with $\sqrt{s} = 13$ TeV constructed with the SM-only hypothesis, while the shaded curves include the effects of four-quark operators with Wilson coefficients $c_i = 10$ TeV^{-2} for illustration. Details of the top quark reconstruction are described in the text.*

at interference level through four linear combinations $\mathcal{O}_{u,d}^{1,2}$. To keep this chapter self-contained, these operators are again displayed in Eq. (4.1).

$$\mathcal{L}_{t\bar{t}} \supset \frac{c_{uG}}{\Lambda^2}(\bar{Q}\sigma^{\mu\nu}T^A u)\tilde{\varphi}G_{\mu\nu}^A + \frac{c_G}{\Lambda^2}f_{ABC}G_\mu^{A\nu}G_\nu^{B\lambda}G_\lambda^{C\mu} + \frac{c_{\varphi G}}{\Lambda^2}(\varphi^\dagger\varphi)G_{\mu\nu}^A G^{A\mu\nu}$$
$$+ \frac{c_{qq}^1}{\Lambda^2}(\bar{Q}\gamma_\mu Q)(\bar{Q}\gamma^\mu Q) + \frac{c_{qq}^3}{\Lambda^2}(\bar{Q}\gamma_\mu\tau^I Q)(\bar{Q}\gamma^\mu\tau^I Q) + \frac{c_{uu}}{\Lambda^2}(\bar{u}\gamma_\mu u)(\bar{u}\gamma^\mu u) \tag{4.1}$$
$$+ \frac{c_{qu}^8}{\Lambda^2}(\bar{Q}\gamma_\mu T^A Q)(\bar{u}\gamma^\mu T^A u) + \frac{c_{qd}^8}{\Lambda^2}(\bar{Q}\gamma_\mu T^A Q)(\bar{d}\gamma^\mu T^A d) + \frac{c_{ud}^8}{\Lambda^2}(\bar{u}\gamma_\mu T^A u)(\bar{d}\gamma^\mu T^A d) \, ,$$

and the four linear combinations of operators are

$$\begin{aligned}
\mathcal{O}_u^1 &= \mathcal{O}_{qq}^{1,1331} + \mathcal{O}_{uu}^{1331} + \mathcal{O}_{qq}^{3,1331} \\
\mathcal{O}_u^2 &= \mathcal{O}_{qu}^{8,1133} + \mathcal{O}_{qu}^{8,3311} \\
\mathcal{O}_d^1 &= \mathcal{O}_{qq}^{3,1133} + \tfrac{1}{4}\mathcal{O}_{ud}^{8,3311} \\
\mathcal{O}_d^2 &= \mathcal{O}_{qu}^{8,1133} + \mathcal{O}_{qd}^{8,3311} \, .
\end{aligned} \tag{4.2}$$

To emphasise that the effects of these operators are most pronounced at high p_T, in Fig. 4.1

we plot the p_T distribution of the hadronic top quark candidate (reconstructed as detailed below) in the Standard Model and for the four quark operator coefficients switched on to a (huge) value of 10 TeV^{-2}, showing the enhancement in the tail.

4.2.1   Analysis details

The events generated from MADEVENT which sample the Wilson coefficient space are subsequently showered by HERWIG++ [350, 351], which takes into account initial and final state radiation showering, as well as hadronisation and the underlying event. At this stage, all our predictions are at leading order in the Standard Model EFT. While considerable progress has recently been made in extending the effective Standard Model description of top quark physics to next-to-leading order [352, 353], the full description of top quark pair production is incomplete at this order. As in chapter 3, we take into account higher-order QCD corrections by re-weighting the Standard Model piece of our distributions to the NLO QCD prediction with K-factors, as obtained from MCFM [107] and cross-checked with MC@NLO [101]. Recently, full NNLO results for top quark pair production have become available in [104, 110, 354], we will comment on their potential for improving our results in Sec. 4.2.2.

We estimate scale uncertainties in the usual way: For the central value of the distributions we choose renormalisation and factorisation scales equal to the top quark mass $\mu_R = \mu_F = m_t$. Then we vary the scales independently over the range $m_t/2 < \mu_{R,F} < 2m_t$. PDF uncertainties are estimated by generating theory observables with the CT14 [92], MMHT14 [93] and NNPDF3.0 [94] as per the recommendations of the PDF4LHC working group for LHC run 2 [308], and we take the full scale+PDF envelope as our theory band. This defines an uncertainty on the differential K-factor which we propagate into each observable. We treat theory uncertainties as uncorrelated with experimental systematics and take them to be fixed as a function of luminosity unless stated otherwise.

Our study focuses on $t\bar{t}$ production at the LHC with $\sqrt{s} = 13$ TeV. For simplicity we focus on the semileptonic decay channel $pp \to t\bar{t} \to q\bar{q}'bl\nu_l b$, where $l \in \{e, \mu\}$ and $q \in \{u, d, s\}$, which strikes the best balance between the large rate but involved jet combinatorics of the fully hadronic channel and the clean signal but low rate and two-neutrino ambiguity of the dilepton case. The sting in the tail for analyses selecting high p_T objects is, of course, low rates. In $t\bar{t}$ production, for instance, only 15% of the cross-section comes from the region $p_T \gtrsim 200$ GeV. We thus aim to quantify at what stage of the LHC programme, if any, the increased sensitivity in the boosted selection can overcome the poorer statistics relative to the resolved selection. We thus construct an analysis which targets both regions simultaneously. Our analysis setup, as implemented in RIVET [355], is as follows (also shown in Tab. 4.1).

Leptons	$p_T > 30$ GeV
	$\|\eta\| < 4.2$
Missing energy	$E_T^{\mathrm{miss}} > 30$ GeV
Small jets	anti-k_T $R = 0.4$
	$p_T > 30$ GeV , $\|\eta\| < 2$
Fat jets	anti-k_T $R = 1.2$
	$p_T > 200$ GeV , $\|\eta\| < 2$
Resolved	≥ 4 small jets w/≥ 2 b-tags
Boosted	≥ 1 fat jet, ≥ 1 small jet w/ b-tag

Table 4.1: *Summary of the physics object definitions and event selection criteria in our hadron-level analysis.*

Firstly, we require a single charged lepton with $p_T > 30$ GeV[†], and find the E_T^{miss} vector as the negative vector sum of the reconstructed momenta, which we require to have a magnitude > 30 GeV. The leptonic W-boson is reconstructed from these by assuming it was produced on-shell. Final state hadrons[‡] are then clustered into jets using the anti-k_T algorithm [356] implemented in FASTJET [357] in two separate groups with $R = (0.4, 1.2)$ requiring $p_T > (30, 200)$ GeV respectively, and jets which overlap with the charged lepton within $\Delta R = 0.3$ are removed. The constituents of the $R = 1.2$ fat jets are reclustered with the Cambridge-Aachen algorithm [358, 359], with all fat jets required to be within $\|\eta\| < 2$, and the $R = 0.4$ small jets are b-tagged within the same η range with an efficiency of 70% and fake rate of 1% [360].

If at least one fat jet and one b-tagged small jet which does not overlap with the leading fat jet exists, we perform a boosted top-tag of the leading fat jet using the HEPTOPTAG-GER [343, 344, 361] algorithm.

The HEPTOPTAGGER procedure is a multistep algorithm optimised to isolate the characteristic three-prong pattern of a hadronically decaying top quark ($t \to Wb \to q\bar{q}'b$). It can be used efficiently for top quark p_T as low as 200 GeV, provided the radius of the large-R jet is large enough to capture all the decay products. Beginning with a Cambridge-Aachen jet J of radius R_{fat} we undo the last step of the clustering, giving two subjets j_1 and j_2 (defined by $m_{j_1} > m_{j_2}$). A *mass-drop* criterion is applied on the heavier subjet:

$$m_{j_1}/m_J < \mu_{\mathrm{frac}}, \tag{4.3}$$

where μ_{frac} is a tuneable parameter. If this criterion is not met the jet is discarded. If it is met, the criterion is applied iteratively on both subjets until all subjets either have

---

[†]We do not consider τ decays here to avoid the more involved reconstruction.

[‡]We do not consider a detector simulation, and B-hadrons are kept stable.

masses less than some input parameter m_{cut} or the jet constituents (tracks, calorimeter deposits) are reached, in which case further unclustering is impossible. If at the end of this unclustering stage, there are less than three subjets, the jet J is discarded. Among the subjets, all possible combinations of 3 subjets are formed (these are called *triplets*). The constituents of the subjets in each triplet are reclustered using the C/A algorithm with a size parameter $R_{\text{filt}} = \min[0.3, \Delta R_{j_1,j_2}/2]$, where $\Delta R_{j_1,j_2}$ is the smallest separation between any two subjets in the triplet. Any constituents of the original jet J that are left outside the reclustered triplets are discarded. This procedure is generically referred to as *filtering*.

All triplets with mass outside the range $140\text{ GeV} \leq m_j \leq 200\text{ GeV}$ are rejected. If more than one is inside the range, the one closest to the top quark mass is selected, this triplet (with 3 or more subjets) is referred to as the top quark candidate. The N_{subjet} highest-p_T subjets of this triplet are chosen. From these subjets, exactly 3 jets are constructed using the C/A algorithm with distance parameter R_{jet} on their constituents. Finally, invariant mass and geometrical requirements on these 3 subjets are applied, to isolate the presence of a W boson decay, namely:

$$R_- < \frac{m_{23}}{m_{123}} < R_+ \quad \text{and} \quad 0.2 < \arctan\frac{m_{13}}{m_{12}} < 1.3$$

$$R_-^2\left(1 + \left(\frac{m_{13}}{m_{12}}\right)^2\right) < 1 - \left(\frac{m_{23}}{m_{123}}\right)^2 < R_+^2\left(1 + \left(\frac{m_{13}}{m_{12}}\right)^2\right) \quad \text{and} \quad \frac{m_{23}}{m_{123}} > 0.35 \tag{4.4}$$

$$R_-^2\left(1 + \left(\frac{m_{12}}{m_{13}}\right)^2\right) < 1 - \left(\frac{m_{23}}{m_{123}}\right)^2 < R_+^2\left(1 + \left(\frac{m_{12}}{m_{13}}\right)^2\right) \quad \text{and} \quad \frac{m_{23}}{m_{123}} > 0.35,$$

where $R_\pm = (1 \pm f_W)(m_W/m_t)$ and f_W is a tuneable parameter of the algorithm between 0 and 1. If at least one of the criteria in Eq. (4.4) are met, the top quark candidate is considered 'tagged'. The HEPTopTagger algorithm can achieve stable efficiencies of around 30% with background contamination of 1% for top candidate p_T ranges from 200 GeV to over 1 TeV (this is the p_T range we consider). To summarise then, the tuneable parameters of the HEPTOPTAGGER algorithm and their optimal values for our analysis are:

$$\{\mu_{\text{frac}}, m_{\text{cut}}, N_{\text{subjet}}, R_{\text{jet}}, f_W\} = \{0.8, 30\text{ GeV}, 5, 0.3, 0.15\}. \tag{4.5}$$

The leptonic top candidate is reconstructed using the leading, non-overlapping (we require $\Delta R(l,j) > 0.4$ for all jets) b-tagged small jet and the reconstructed leptonic W. If no fat jet fulfilling all the criteria exists, we instead require at least 2 b-tagged small jets and 2 light small jets. If these exist we perform a resolved analysis by reconstructing the hadronic W-boson by finding the light small jet pair that best reconstructs the W mass, and reconstruct the top candidates by similarly finding the pairs of reconstructed W-bosons and b-tagged small jets that best reconstruct the top mass.
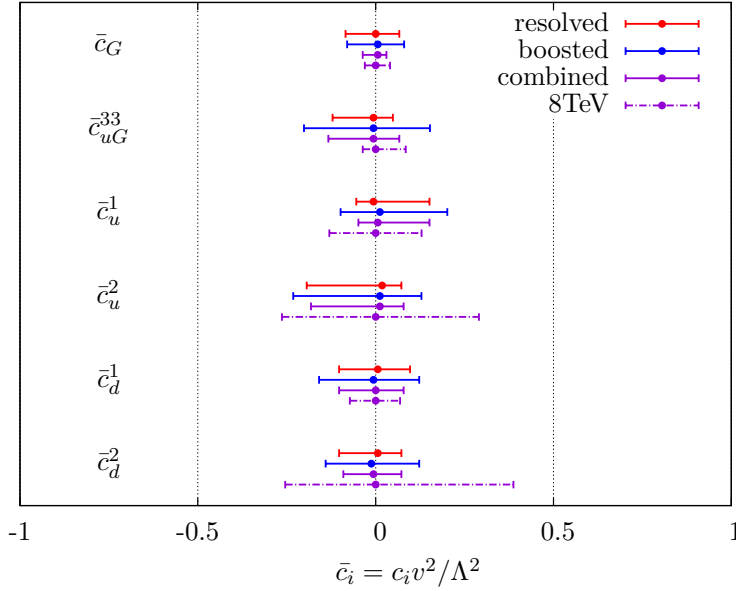
Figure 4.2: *Individual 95% bounds on the operators considered here, from the boosted analysis and the resolved fat jet analysis, and the combined constraint from both, assuming 20% systematics and 30 fb^{-1} of data. We also show existing constraints from unfolded 8 TeV p_T distributions published in [292] and [362], showing the sizeable improvement even for a modest luminosity gain.*

Finally, regardless of the approach used, we require both top candidates to have $|m_{\text{cand}} - m_{\text{top}}| < 40$ GeV. If this requirement is fulfilled the event passes the analysis.

4.2.2 Results

Impact of experimental precision

Using a sample size of 30 fb^{-1} with a flat 20% systematic uncertainty (motivated by typical estimates from existing experimental analyses by ATLAS [363] and CMS [364]) on both selections as a first benchmark, and the p_T distribution of Fig. 4.1, the 1-dimensional 95% confidence intervals on the operators considered here are presented in Fig. 4.2. All the bounds presented here are 'one-at-a-time', i.e. we do not marginalise over the full operator set. Our purpose here is to highlight the relative contributions to the allowed confidence intervals here, rather than to present a global operator analysis.

As a general rule, the increased sensitivity to the Wilson coefficients offered by the boosted selection is overpowered by the large experimental systematic uncertainties in this region, and the combined limits are dominated by the resolved top quarks. The exception to this rule is the coefficient c_G from the operator $\mathcal{O}_G = f_{ABC} G_\nu^{\mu,A} G_\lambda^{\nu,B} G_\mu^{\lambda,C}$. Expanding out the field strength tensors leads to vertices with up to six powers of momentum in the
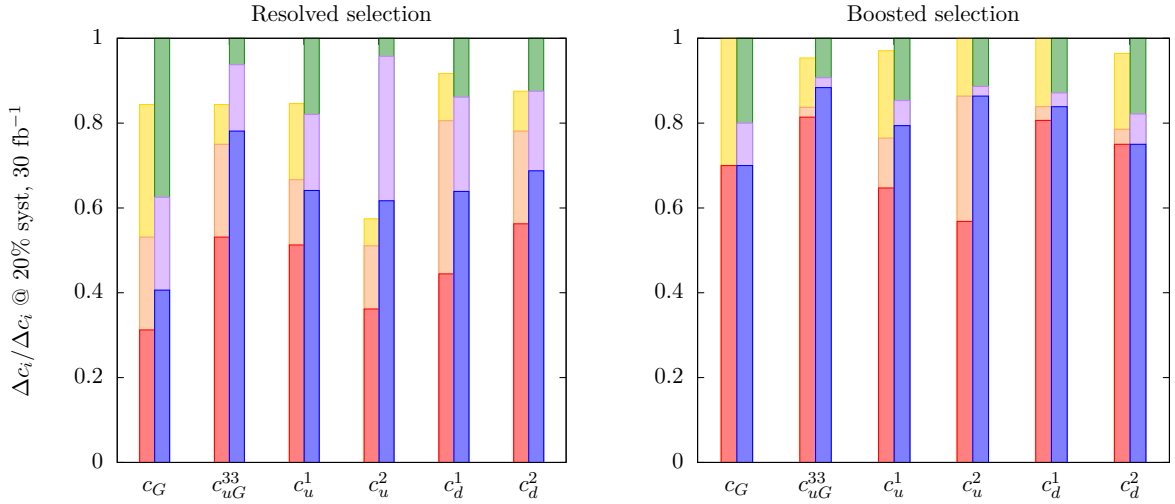
Figure 4.3: *Fractional improvement on the 95% confidence intervals for the operators considered here, with various combinations of luminosity and experimental systematics considered. We take the width of the 95% confidence limit obtained from 20 % systematic uncertainty and 30 fb^{-1} of data as a baseline (green bar), and normalise to this, i.e. we express constraints as a fractional improvement on this benchmark. The purple and blue bars represent respectively, 300 fb^{-1} and 3 ab^{-1} of data, also at 20% systematics, while the yellow, orange and red are the analogous data sample sizes for 10% systematics.*

numerator, more than enough to overcome the naïve $1/\hat{s}^2$ unitarity suppression. Large momentum transfer final states thus give stronger bounds on this coefficient, even with comparatively fewer events.

With these constraints as a baseline, it is then natural to ask by how much they can be improved upon when refinements to experimental precision are made. The constraints are presented in Fig. 4.3 for different combinations of systematic and statistical uncertainties. We take the width of the 95% confidence interval in Fig. 4.2 as our normalisation (the green bars), and express the fractional improvements on the limits that can be achieved relative to this baseline, for each operator. The right bars (green, purple, blue) represent 20% systematic uncertainties with, respectively 30, 300 and 3 ab^{-1} of data. The left bars (yellow, orange, red) represent the same respective data sample sizes, but with 10% systematic uncertainties.

Beginning with the resolved selection, we find that the limits on the coefficient c_G can be improved by 40% by going from 30 fb^{-1} to 300 fb^{-1}, and by a further 20% when the full LHC projected data sample is collected. Systematic uncertainties have a more modest effect on this operator: at 3 ab^{-1} the limit on c_G is only marginally improved by a 10% reduction in systematic uncertainty. This merely reflects that c_G mostly impacts the high p_T tail, so it can only be improved upon in the threshold region by collecting enough
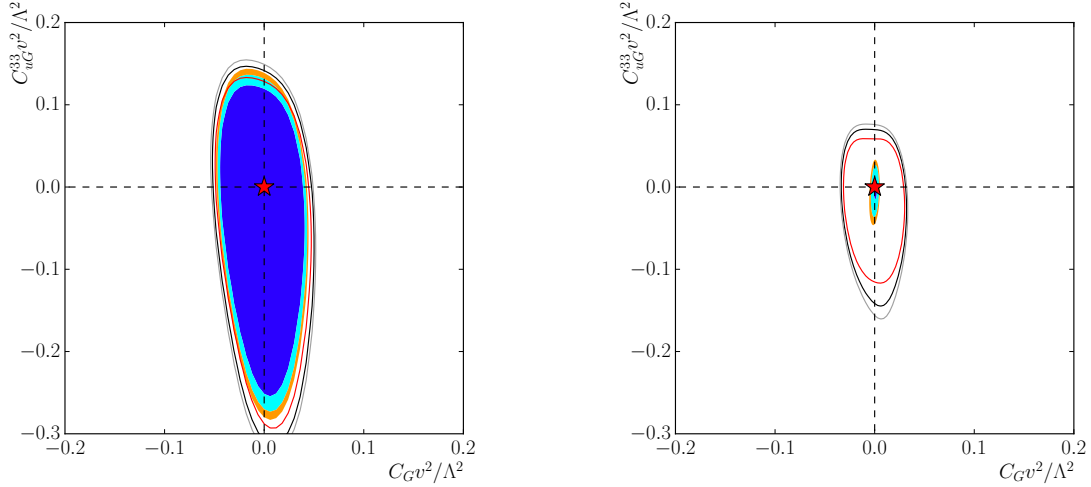
Figure 4.4: *Left: 68%, 95% and 99% confidence intervals for c_G and c_{uG}^{33}, the lines are obtained using experimental (20% systematics and 30 fb^{-1} of data) uncertainties along with theoretical uncertainties, the filled contours using only experimental uncertainties. Right: the same plot, but using 10% systematics and 3 ab^{-1} of data, showing the much stronger impact of theory uncertainties in this region.*

data to overcome the lack of sensitivity. 8 TeV measurements are already constraining the relevant phase space region efficiently and the expected improvement at 13 TeV is only mild (see below).

For the chromomagnetic dipole operator $\mathcal{O}_{uG}^{33}$, improving the experimental systematics plays much more of a role. A 10% improvement in systematics, coupled with an increase in statistics from 30 fb^{-1} to 300 fb^{-1} leads to stronger limits that maintaining current systematics and collecting a full 3 ab^{-1} of data. Similar conclusions apply for the four-quark operators, to varying degrees, i.e. reducing systematic uncertainties can provide comparable improvements to collecting much larger data samples.

For the boosted selection, the situation is quite different. For all the operators we consider, improving systematic uncertainties by 10% has virtually no effect on the improvement in the limits. This simply indicates that statistical uncertainties dominate the boosted region at 30 fb^{-1}. For c_G, at 300 fb^{-1} some improvement can be made if systematics are reduced, however we then see that systematic uncertainties saturate the sensitivity to c_G, i.e. there is no improvement to be made by collecting more data. For c_{uG}^{33}, a modest improvement can also be made both by reducing systematics by 10% and by increasing the dataset to 300 fb^{-1}. However, going beyond this, the improvement is minute. The four-quark operators again follow this trend, although c_u^2 shows much more of an improvement when going from 300 fb^{-1} to 3 ab^{-1}.

The role of theory uncertainties

The other key factor in the strength of our constraints is the uncertainties that arise from theoretical modelling. The scale and PDF variation procedure typically leads to uncertainties in the 10-15% range. Fully differential K-factors for top pair production at NNLO QCD (i.e. to order $\mathcal{O}(\alpha_s^4)$) have become available, which have substantially reduced the scale uncertainties. The numbers quoted in Refs. [109, 110] are for the Tevatron and 8 TeV LHC, and available only for the low to intermediate p_T^t range ($p_T^t < 400$ GeV). Updated results for 13 TeV have become available only recently [365]. It is worthwhile to ask what impact such an improvement could have on the constraints.

We put this question on a firm footing by showing in Fig. 4.4 the 2D exclusion contours for the coefficients c_G and c_{uG}^{33}, as obtained from combining the boosted and resolved limits, at fixed luminosity and experimental systematics, first using our NLO theory uncertainty, and also using *no theory uncertainty at all*. For 30 fb^{-1} the improvement is limited, indicating that at this stage in the LHC programme the main goal should be to first improve experimental reconstruction of the top quark pair final state. However, at 3 ab^{-1} the improvement is substantial, indicating that it will also become necessary to improve the theoretical modelling of this process, if the LHC is to augment its kinematic reach for non-resonant new physics.

In addition to SM theoretical uncertainties, there are uncertainties relating to missing higher-order terms in the EFT expansion. Uncertainties due to to loop corrections and renormalisation-group flow of the operators $\mathcal{O}_i^{(6)}$ are important for measurements at LEP-level precision [237, 366] where electroweak effects are also resolved. However, at the LHC we find them to be numerically insignificant compared to the sources of uncertainty that we study in detail here. In addition, there is also the possibility of large effects due to $D = 8$ operators, particularly owing to additional derivatives in the EFT expansion. Since the interference effects of omitted $D = 8$ operators are formally of the same order as the retained quadratic terms in the $D = 6$ operators, we emphasise that the numerical constraints presented here should be treated with caution. The only way to be certain that the omission of these terms is justified is to compute the effects of the interference of the relevant $D = 8$ operators to a given process and demonstrate them to be small. This has been shown to be true for the $gg \to t\bar{t}$ subprocess [215, 216]. However, due to the large number of operators present there, this has not been studied for the $q\bar{q} \to t\bar{t}$ process. We leave a full computation of these effects as a future direction of study.

4.2.3 Interpreting the results

The whole purpose of the EFT approach is to serve as a bridge between the Standard Model and heavy degrees of freedom residing at some unknown mass scale M_*. Connect-
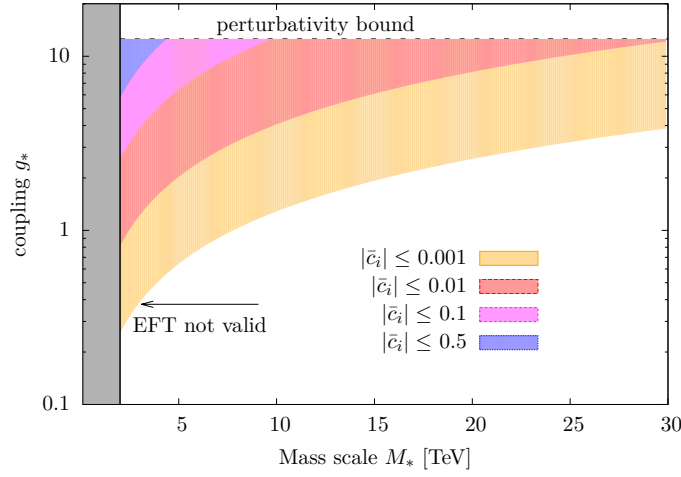
Figure 4.5: *Areas in the new coupling-BSM mass scale plane (see also [320]), resulting from our fit coverage. Shaded areas are constrained in perturbative UV completions at a scale M_*, subject to the boundary condition Eq. (4.6). The shaded grey area is probed by the pseudodata of our fit. We do not consider unitarity bounds in this work.*

ing the EFT to this scale, however, necessarily involves making assumptions about the couplings of this new physics. We can make statements about the relation between the constraints presented here and such a scale, however, by making general assumptions, such as perturbativity of the underlying new physics.

Consider, for example, the simple case where the perturbative UV physics is characterised entirely by a single coupling g_* and a unique mass scale M_*. Such a scenario could arise from integrating out a heavy, narrow resonance. In this case we have the simple tree-level matching condition

$$\frac{c_i}{\Lambda^2} = \frac{g_*^2}{M_*^2}. \tag{4.6}$$

Constraints on c_i then map onto allowed regions in the g_*-M_* plane. In Fig. 4.5 we sketch these regions for illustrative values of c_i. In order for the EFT description of a given mass region to be valid, we must not resolve it our measurement. Therefore we impose a hard cut at $\sqrt{s} = 2$ TeV, obtained from the maximum $t\bar{t}$ invariant mass probed in our SM pseudodata. We also impose a generic perturbativity restriction $g_* \lesssim 4\pi$ to ensure that our EFT expansion is well-behaved and higher-dimensional operators do not affect the power counting.

We see that for large Wilson coefficients $\bar{c}_i \gtrsim 0.5$ only a very small window of parameter space may be constrained, but the weak limits push the underlying coupling to such large values that loop corrections are likely to invalidate the simple relation of Eq. (4.6), making it hard to trust these limits. However, at 3 ab^{-1}, the projected constraints are typically $\bar{c}_i \lesssim 0.01$, therefore, even for moderate values of the coupling g_*, our constraints are able

to indirectly probe mass scales much higher than the kinematic reach of the LHC.

4.2.4 Discussion

The special role of the top quark in BSM scenarios highlights the importance of searches for new interactions in the top sector. Taking the lack of evidence of resonant new physics in the top sector at face value [367–369], we can assume that new interactions are suppressed by either weak couplings or large new physics scales. In both cases we can analyse the presence of new physics using effective field theory techniques. A crucial question that remains after the results from the LHC run 1 is in how far a global fit from direct search results will improve with higher statistics and larger kinematic coverage. We address this question focusing on the most abundant top physics-related channel $pp \to t\bar{t}$, which probes a relevant subset of top quark effective interactions. In particular, we focus on complementary techniques of fully-resolved vs. boosted techniques using jet-substructure technology, which are affected by different experimental systematic uncertainties. Sensitivity to new physics is a trade off between small statistical uncertainty and systematic control for low p_T final states at small new physics-induced deviations from the SM expectation (tackled in fully-resolved analyses) and the qualitatively opposite situation at large p_T. For the typical parameter choices where top-tagging becomes relevant and including the corresponding efficiencies, we can draw the following conclusions:

- Boosted top kinematics provide a sensitive probe of new interactions in $t\bar{t}$ production mediated by modified trilinear gluon couplings. In particular, this observation shows how differential distributions help in breaking degenerate directions in a global fit by capturing sensitivity in phenomenologically complementary phase space regions.

- The sensitivity to all other operators detailed in Eq. (4.1) is quantitatively identical for boosted and fully-resolved analyses for our choice of $p_T^{\text{boost}} \geq 200$ GeV. Increasing the boosted selection to higher p_T (where the top tagging will become more efficient) will quickly move sensitivity to new physics effects to the fully resolved part of the selection. The boosted selection is saturated by large statistical uncertainties for the for the typical run 2 luminosity expectation. These render systematic improvements of the boosted selection less important in comparison to the fully resolved selection, which provides an avenue to set most stringent constraint from improved experimental systematics. Similar observations have been made for boosted Higgs final states [370] and are supported by the fact that the overflow bins in run 1 analyses provide little statistical pull [2].

- Theoretical uncertainties that are inherent to our approach are not the limiting factors of the described analysis in the forseeable future, but will become relevant when

127

statistical uncertainties become negligible at very large integrated luminosity.

Boosted analyses are highly efficient tools in searches for resonant new physics [367–369, 371]. Our results show that similar conclusions do not hold for non-resonant new physics effects when the degrees of freedom in question do not fall inside the kinematic coverage of the boosted selection anymore. Under these circumstances, medium p_T range configurations which maximise new physics deviation relative to statistical and experimental as well as theoretical uncertainty are the driving force in setting limits on operators whose effects are dominated by interference with the SM amplitude in the top sector. This also implies that giving up the boosted analysis in favor of a fully resolved analysis extending beyond $p_T^t \geq 200$ GeV will not improve our results significantly. The relevant phase space region can be accessed with fully resolved techniques, with a large potential for improvement from the experimental systematics point of view.

4.3   Associated Z production projections

We have seen that there are good improvement prospects for the $t\bar{t}$ $D = 6$ operators. Of the constraints listed in chapter 3, however, by far the weakest are those extracted from top quark neutral couplings in $t\bar{t}Z$ and $t\bar{t}\gamma$ production. This is because these processes have a much smaller rate, so at this stage of the LHC programme their measurements are currently statistics dominated. It is also natural to then ask how they may be improved over the lifetime of the LHC. This is the subject of this section.

4.3.1   Top electroweak couplings

In the SM, the electroweak $t\bar{t}Z$ coupling is given by the vector-axial-vector coupling

$$\mathcal{L}_{\text{ttZ}} = e\bar{t} \left[\gamma^\mu (v_t - \gamma_5 a_t) \right] t Z_\mu \tag{4.7}$$

where

$$\begin{aligned}
v_t &= \frac{T_t^3 - 2Q_t \sin^2 \theta_W}{2 \sin \theta_W \cos \theta_W} \simeq 0.24, \\
a_t &= \frac{T_t^3}{2 \sin \theta_W \cos \theta_W} \simeq 0.60.
\end{aligned} \tag{4.8}$$

To capture effects beyond the SM in this Lagrangian there are two approaches: one can write down anomalous couplings for the $t\bar{t}Z$ vertex, such that $\mathcal{L}_{\text{ttZ}}$ receives a term

$$\Delta\mathcal{L}_{\text{ttZ}} = e\bar{t} \left[\gamma^\mu (C_{1V} + \gamma_5 C_{1A}) + \frac{i\sigma^{\mu\nu} q_\nu}{2M_Z} (C_{2V} + \gamma_5 C_{2A}) \right] t Z_\mu, \tag{4.9}$$

where $q = p_t - p_{\bar{t}}$. While this has the advantage of elucidating the various spin structures that can impact the $t\bar{t}Z$ vertex, it has the drawback that it does not allow for a simple power counting of which anomalous couplings would have the strongest effect. For example, the coefficient C_{2A} is zero in the Standard Model, so that any corrections to it come solely from new physics contributions, which should be smaller than couplings that have SM interference.

To augment this description, one can instead supplement Eq. (4.7) with higher-dimensional operators. At leading order in the SMEFT, the list of operators that generate modifications to the $t\bar{t}Z$ vertex is, expressed in the basis and notation of Ref. [199]:

$$
\begin{aligned}
\mathcal{O}_{uW} &= (\bar{Q}\sigma^{\mu\nu}u)\tau^I\tilde{\varphi}W^I_{\mu\nu} \\
\mathcal{O}_{uB} &= (\bar{Q}\sigma^{\mu\nu}u)\tilde{\varphi}B_{\mu\nu} \\
\mathcal{O}^{(3)}_{\varphi q} &= (\varphi^\dagger i\overleftrightarrow{D}^I_\mu\varphi)(\bar{Q}\tau^I\gamma^\mu Q) \\
\mathcal{O}^{(1)}_{\varphi q} &= (\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar{Q}\gamma^\mu Q) \\
\mathcal{O}_{\varphi u} &= (\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar{u}\gamma^\mu u).
\end{aligned}
\tag{4.10}
$$

The dictionary between the $D = 6$ operators of Eq. (4.10) and the anomalous couplings of Eq. (4.9) is

$$
\begin{aligned}
C_{1V} &= \frac{v^2}{\Lambda^2}\Re\left[c^{(3)}_{\varphi q} - c^{(1)}_{\varphi q} - c_{\varphi u}\right]^{33} \\
C_{1A} &= \frac{v^2}{\Lambda^2}\Re\left[c^{(3)}_{\varphi q} - c^{(1)}_{\varphi q} + c_{\varphi u}\right]^{33} \\
C_{2V} &= \sqrt{2}\frac{v^2}{\Lambda^2}\Re\left[\cos\theta_W c_{uW} - \sin\theta_W c_{uB}\right]^{33} \\
C_{2A} &= \sqrt{2}\frac{v^2}{\Lambda^2}\Im\left[\cos\theta_W c_{uW} + \sin\theta_W c_{uB}\right]^{33},
\end{aligned}
\tag{4.11}
$$

where the superscript 33 denotes that we are considering the 3rd generation only in the fermion bilinears of Eq. (4.10). Since $c^{(3)}_{\varphi q}$ and $c^{(1)}_{\varphi q}$ only appear with an overall opposite sign, we can only constrain the operator $\mathcal{O}^{(3)}_{\varphi q} - \mathcal{O}^{(1)}_{\varphi q} \equiv \mathcal{O}_{\varphi q}$ from $t\bar{t}Z$ couplings. We will discuss a method for bounding the two operators independently later in the chapter.

C_{2A} is generated by a $\mathcal{CP}$-odd combination of operators, therefore it does not interfere with SM amplitudes and so its effects are expected to be smaller. Since in this study we are more interested in the absolute mass scales of these operators, we set all Wilson coefficients to be real, however we note that $\mathcal{CP}$-sensitive observables such as angular distributions can also distinguish the $\mathcal{CP}$ character of the Wilson coefficients. We also assume that the new physics solely impacts the $t\bar{t}Z$ vertex, so we do not consider operators which modify the $Ze\bar{e}$ vertex, nor four-fermion operators which can contribute to the $q\bar{q} \to t\bar{t}$ or $e^+e^- \to t\bar{t}$ processes (see e.g. Refs. [2, 372] for constraints on the former).
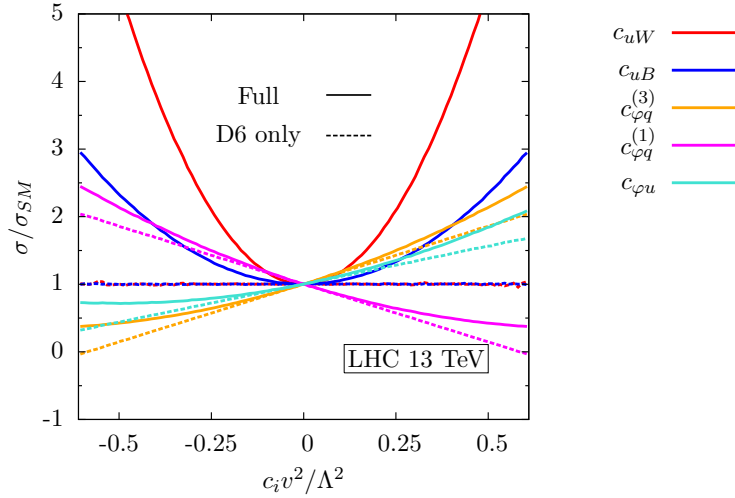
Figure 4.6: *Ratio of the full SM $pp \to t\bar{t}Z$ cross-section with the operators of Eq. (4.10) switched on individually to the NLO Standard Model estimate. The dashed lines show the contribution from the interference term, and the solid lines show the full dependence.*

4.3.2 Total rates

To appreciate the impact of the operators of Eq. (4.10), in Fig. 4.6 we plot the ratio of the full $t\bar{t}Z$ cross-section with each operator switched on individually, to the NLO SM prediction, taken from Ref. [373]. For ease of interpretation, we split up the cross-section into the contribution from the interference term and the quadratic term. We see firstly that the operators $\mathcal{O}_{uW}$ and $\mathcal{O}_{uB}$ have the strongest impact on the total cross-section, but this comes purely from the squared term (this was also noted in Ref. [374]). The remaining operators have a milder effect on the cross-section, but their interference term dominates. We also see that the operators $\mathcal{O}_{\varphi q}^{(3)}$ and $\mathcal{O}_{\varphi q}^{(1)}$ contribute the same dependence but with an opposite sign, as discussed in Sec. 4.3.1, therefore we can only bound the linear combination $\mathcal{O}_{\varphi q}$.

The LHC bounds on the coefficients of these operators from 8 TeV $t\bar{t}Z$ production cross-sections were presented in chapter 3. The current constraints are weak. Since then, ATLAS and CMS have presented measurements using 13 TeV collision data, with measured values 0.9 ± 0.3 pb [375] and 0.7 ± 0.21 pb [376], respectively. The constraints on the operators using these two measurements are shown in Fig. 4.7, where the coefficients are normalised to the 'bar' notation $\bar{c}_i = c_i v^2/\Lambda^2$, and the operators are switched on individually.

We see that the current constraints are still quite weak, mainly due to the large ($\sim$ 30%) experimental uncertainties. These measurements are currently statistics dominated, so it is instructive to ask what the expected improvement is over the lifetime of the LHC. Using a constant systematic uncertainty of 10% based on the current estimate, we also
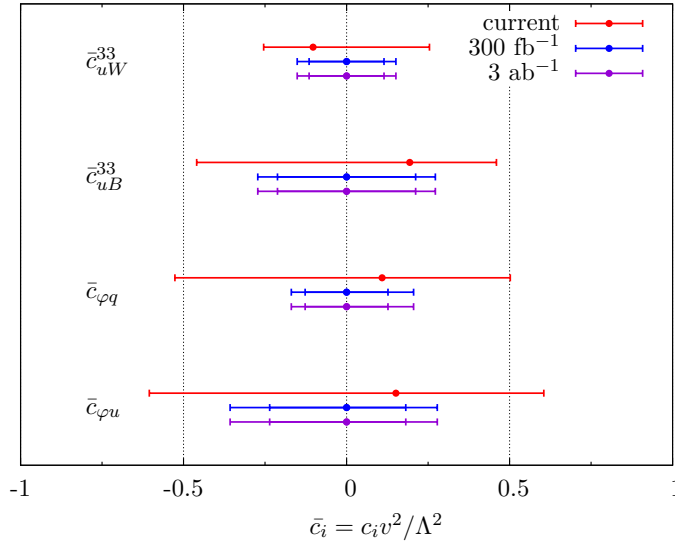
Figure 4.7: *Individual 95% confidence intervals on the coefficients of the operators of Eq. (4.10) using the current 13 TeV measurements (red bars). Also shown are the projected constraints using 300 fb^{-1} (blue) and 3 ab^{-1} (purple) of SM pseudodata. For the latter two cases, the inner bars show the improvement when theory uncertainties are reduced to 1%.*

plot in Fig. 4.7 the constraints using 300 fb^{-1} and 3 ab^{-1} of SM pseudodata. We see that there will be an improvement by factors of 1.5 to 2 by the end of Run III, but after this the measurement is saturated by systematics.

To highlight the benefits of improving the theory description in tandem, we also show in Fig. 4.7 the projected constraints if theory uncertainties are improved to 1% from the current $\mathcal{O}(10\%)$ precision, which does not seem unreasonable over the timescales we are considering. We see again that there will be no subsequent improvement after 300 fb^{-1} unless experimental systematics are reduced.

4.3.3   Impact of differential distributions

Finally, it should be noted that as more data becomes available, it may be possible to measure $t\bar{t}Z$ cross-sections differentially in final state quantities. Since cuts on the final state phase space can enhance sensitivity to the region where naïve power counting says $D = 6$ operators become more important, differential distributions could substantially improve the fit prospects, as has already been demonstrated for $t\bar{t}$ production. [2,3].

To illustrate this, in Fig. 4.8 we plot the distributions for the Z boson transverse momentum and top pair transverse momentum, both for the SM only case and with each operator switched on to a value of $\bar{c}_i \simeq 0.3$; approximately the maximum allowed by current constraints in Fig. 4.7. We see that extra enhancement in the tail is visible for the field
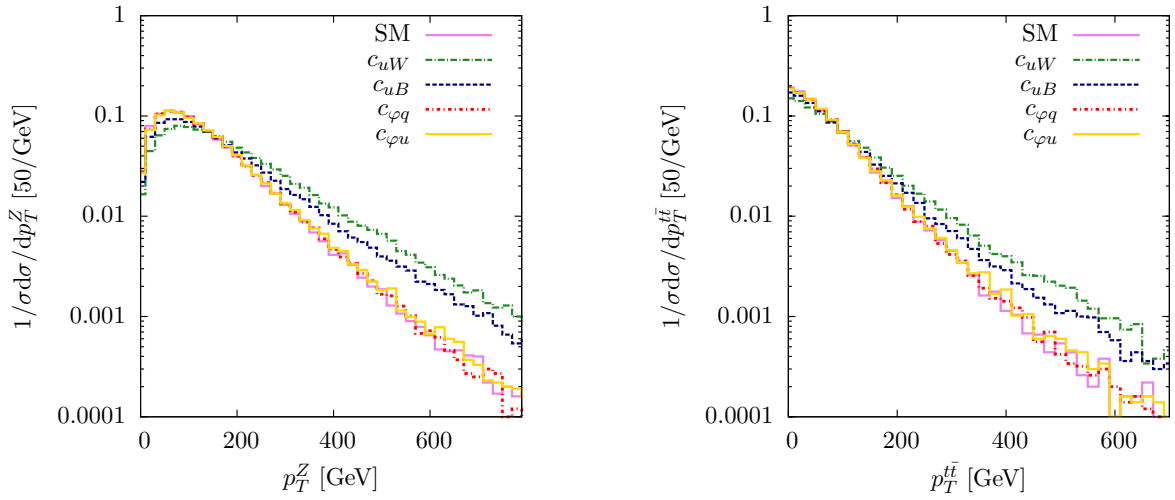
Figure 4.8: *Kinematic distributions in pp → tt̄Z production at 13 TeV for the SM prediction and for the operators of Eq. (4.10) switched on to their maximum value allowed by current data. Left: the Z boson transverse momentum spectrum. Right: $p_T^{t\bar{t}} = p_T^t - p_T^{\bar{t}}$ spectrum. All distributions are normalised to the total cross-section. Shape differences can be seen in the tails for the operators $\mathcal{O}_{uW}$ and $\mathcal{O}_{uB}$, showing that differential distributions provide complementary information to overall rates.*

strength tensor operators $\mathcal{O}_{uW}$ and $\mathcal{O}_{uB}$, due to the extra momentum dependence in the numerator from the field strength tensor. For the φ-type operators, since the interference is solely proportional to $\varphi^\dagger \varphi \to v^2/\Lambda^2$, there is no extra enhancement at high p_T.

We do not estimate the improvement of the fit by taking these distributions into account, since this would require proper estimates of experimental systematics and tracking the nontrivial correlations between the kinematic quantities in the massive 3-body final state. Here, we merely comment that it may be an avenue worth pursuing as more data becomes available.

4.4   Future collider prospects

We see that despite the impressive statistical sample of top quark data that enters these fits, the subsequent direct bounds on the operator coefficients, and, by extension, the scale of new physics that would generate those operators, are rather weak [2]. There are few top quark measurements at the LHC that can be considered "precision observables" (helicity fractions in top decays are an exception [145]). By dimensional analysis, the strength of the interference of these operators with energy $\sqrt{s}$ typically scales as s/Λ^2, and the areas of phase space that are most sensitive are plagued by correlated experimental and theoretical systematics. Moreover, the associated weak limits translate into values of Λ

that are probed by the high energy bins of the measurement, bringing into question the validity of the truncated EFT description [377] and care needs to be taken when combining measurements of different exclusive energy ranges of a binned distribution [320]. Inclusive cross-sections, being typically dominated by the threshold region $\sqrt{s} \sim (2)m_t$, are under more theoretical control, but bring far less sensitivity.

Lepton colliders are not vulnerable to either of these problems. Firstly, there is excellent control over the hard scale of the interaction $\sqrt{s}$, so one can always ensure that the limits on the $D = 6$ operators are consistent with a well-behaved EFT expansion.[§] Secondly, the theoretical uncertainties from Standard Model calculations are much smaller: there are no PDFs, and the current state of the art precision for $t\bar{t}$ production is N^{3}LO QCD at fixed-order [378], and NNLO+NNLL including threshold resummation, which bring SM scale uncertainty variation bands to the percent level [379].

The physics case for a e^+e^- collider is by now well-established. The principal motivation is to perform a detailed precision study of the couplings of the Higgs boson in the much cleaner environment that a lepton collider affords, which will bring Higgs coupling measurements to an accuracy that will not be challenged by the LHC, even after it collects 3 ab^{-1} of data [380]. The electroweak couplings of the top quark are also clearly within the remit of such a collider. Currently, the only handle on top quark electroweak couplings from the LHC is through the associated production $pp \to t\bar{t}V$ where $V \in \{Z, W, \gamma\}$. Whilst measurements of these processes are now approaching the 5σ level, the pull that they have on a global fit is small [2]. Measurements of electroweak single top production bring stronger bounds, but are sensitive to a smaller subset of operators.

At a lepton collider, on the other hand, the process $e^+e^- \to Z^*/\gamma \to t\bar{t}$ is extremely sensitive to top electroweak couplings. While the overall rate is more modest than at the LHC due to the parametric α_{EW}/α_s and s-channel suppression, the process is essentially background-free, and would constitute the first true precision probe of the electroweak sector of the top quark, and open up a new avenue for top quark couplings, complementary to the well-studied top QCD interactions. Several studies of the prospects for improvement of top measurements at future colliders have already been undertaken (see for example Refs. [381–386]), in particular for the proposed International Linear Collider (ILC), but none have explicitly quantified the gain in the constraints on the top electroweak sector of the SMEFT, nor provided a comparative study of different collider options. The remainder of this chapter provides such a study.

Going beyond the LHC, currently, the most mature proposal is for a linear e^+e^- collider with a centre of mass energy ranging from 250 GeV to up to 1 TeV. There are several

---

[§]Consistently improving the perturbative precision within the dimension 6 framework, however, makes the truncation of the perturbative series necessary as corrections to $(D = 6)^2$ operators will typically require unaccounted for $D = 8$ counterterms.
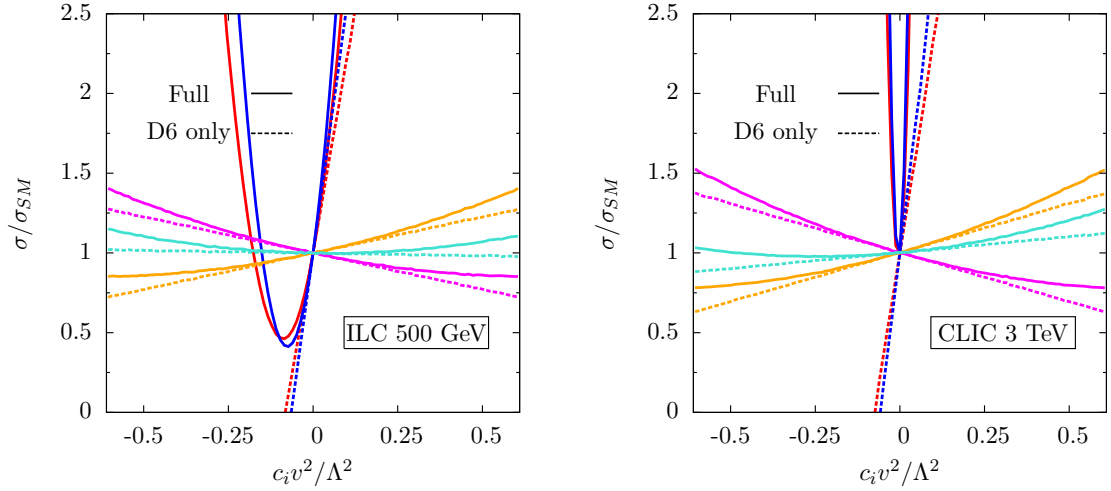
Figure 4.9: *Left: Ratio of the full SM e^+e^- to $t\bar{t}$ cross-section at $\sqrt{s}$ = 500 GeV with the operators of Eq. (4.10) switched on individually to the NLO Standard Model estimate. The dashed lines show the contribution from the interference term, and the solid lines show the full dependence The operator colour-coding is the same as Fig. 4.6. Right: Likewise for CLIC running at $\sqrt{s}$ = 3 TeV.*

scenarios for integrated luminosity and CM energy combinations. The most-studied is the so-called H-20 option, which involves running at 500 GeV for 500 fb^{-1} of data, followed by 200 fb^{-1} of data at the $t\bar{t}$ threshold to perform detailed measurements of the top quark mass, and 300 fb^{-1} of data at $\sqrt{s}$ = 250 GeV to maximise the machine's Higgs potential with high precision. After a luminosity upgrade, a further 3.5 ab^{-1} is gathered at $\sqrt{s}$ = 500 GeV, followed by another $\sqrt{s}$ = 250 GeV run at 1.5 ab^{-1}. Since we are most interested in the ILC mass reach for new physics, in this study we focus on the 500 GeV ILC running.

An important parameter for lepton colliders is the energy spread of electron and positron beams (see e.g. [387]). In order to estimate the effect on our results, we use the results of [387] to calculate the expected change in the cross-section by including the effects of initial state radiation, beam spread and beamstrahlung. We find that for the typical beam profile, the associated uncertainty is not a limiting factor and we neglect these effects in the following.

4.4.1 The $t\bar{t}$ total cross-section

Top pair production has a more modest rate here than at a hadron collider. The state-of-the-art Standard Model calculations for (unpolarised) $e^+e^- \to t\bar{t}$ production are at N^{3}LO QCD [378,379], and at NLO EW [388] (with partial NNLO results in Ref. [389]) and predict a cross-section $\sigma \simeq 0.57$ pb. The conventional scale variation gives a QCD uncertainty at the per-mille level. While this rate is more than a factor of a thousand smaller than at the
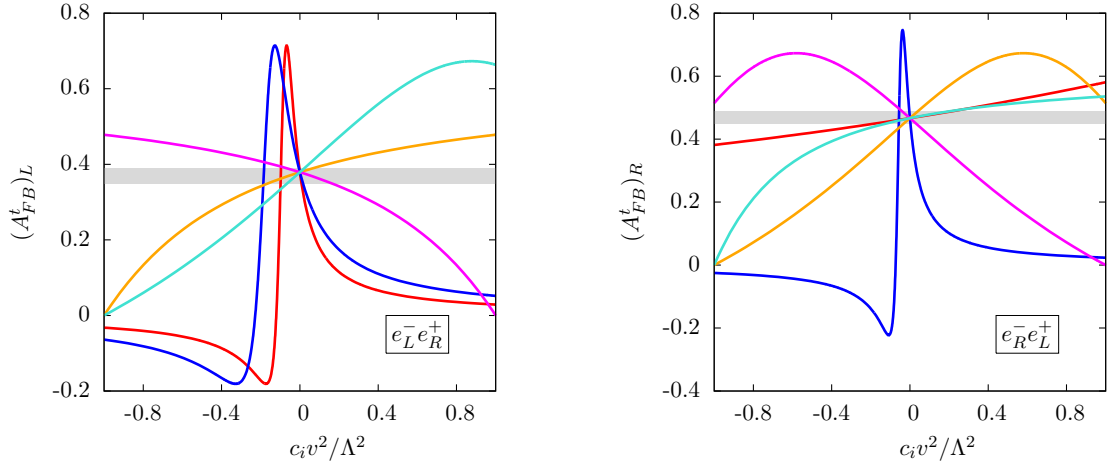
Figure 4.10: *Full dependence of the $t\bar{t}$ forward-backward asymmetry of Eq. (4.12) on the operators of Eq. (4.10) for left-handed polarised electrons (left) and right-handed polarised electrons (right) at $\sqrt{s}$ = 500 GeV, the operator colour coding is the same as Fig. 4.6. We also show a 5% uncertainty band around the SM prediction, to estimate the expected constraints.*

13 TeV LHC, the process is essentially background free. Thus, after even 500 fb^{-1} of data the statistical uncertainty will be approximately 0.2%, and so completely subdominant to the systematics.

We can thus repeat the exercise of extracting the bounds on the coefficients of the operators of Eq. (4.10) using SM pseudodata. As a guide for the expected numerical constraints, we also plot the ratio of the total cross-section in the presence of the operators to the SM prediction, this time using the total (unpolarised) cross-section at the 500 GeV ILC. This is shown on the left of Fig. 4.9.

We see again that the operators $\mathcal{O}_{uW}$ and $\mathcal{O}_{uB}$ are the strongest, however, unlike the case of $t\bar{t}Z$ production the interference term dominates at small c_i/Λ^2. The result of this is that there is a cancellation between the interference and quadratic terms at approximately $c_i/\Lambda^2 \simeq -3$ TeV^{-2}, leading to a SM-like cross-section and a second, degenerate minimum in the χ^2. The constraints obtained from a one-at-a-time fit of these operators to the SM pseudodata is shown in the red bars on the right of Fig. 4.11.

The operators $\mathcal{O}_{uW}$ and $\mathcal{O}_{uB}$ are very tightly constrained, due to their much stronger impact on the cross-section stemming from the extra momentum dependence flowing through the vertex. The φ-type operators are more weakly constrained, but on the whole the constraints are typically 100 times stronger than for the LHC $t\bar{t}Z$ production projections in Sec. 4.3, which is unsurprising giving the difference in precision.

Individual constraints are less useful in practice, however. Firstly, in a plausible UV
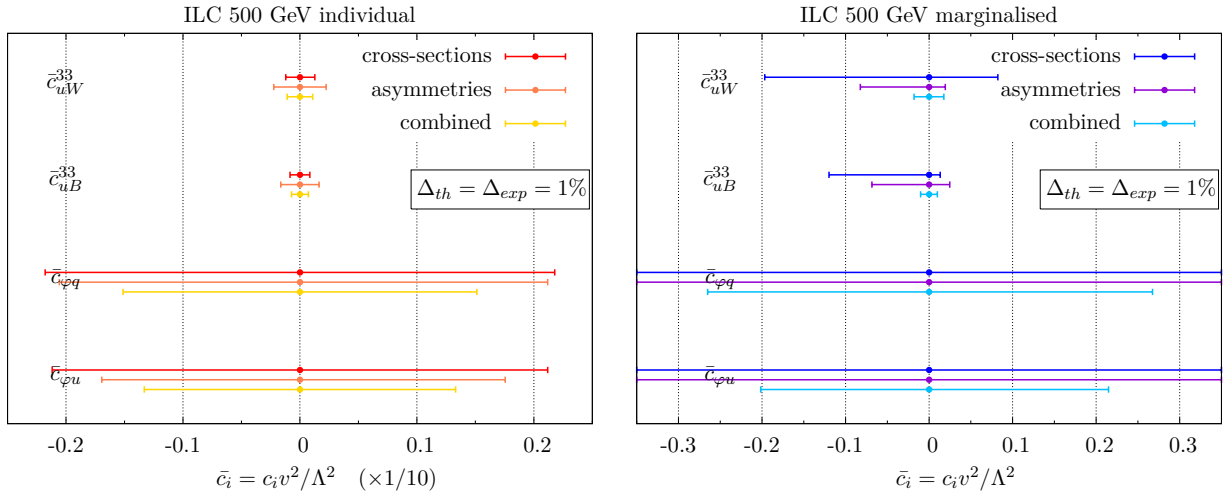
Figure 4.11: *95% confidence ranges for the operators we consider here, from the 500 GeV ILC, assuming 1% theoretical and experimental uncertainties, by fitting to cross-sections, asymmetries, and the combination, with each operator considered individually (left) or in a 5D fit (right). To display both on the same plot, we scale the individual constraints up by a factor of 10, so that the bottom axis is actually $\bar{c}_i/10$.*

scenario that would generate these operators one would typically expect more than one to be generated at once, so that one-at-a-time constraints cannot be straightforwardly linked to a specific 'top-down' model. Secondly, there can in general be cancellations between different operators for a given observable that can yield spurious local minima and disrupt the fit. This would not be visible in the individual constraints, and so would obscure degeneracies in the operator set that could be broken by considering different observables. Therefore, we also consider constraints where we marginalise over the remaining three coefficients in the fit, as also discussed in chapter 3. These are shown in the blue bars on the right of Fig. 4.11.[¶]

We see that, with the exception of $\mathcal{O}_{uW}$ and $\mathcal{O}_{uB}$, marginalising over the full operator set wipes out the constraints. This is because even for large values of coefficients, the pull that a particular operator has on the cross-section can easily be cancelled by another operator. We can conclude that, despite the impressive precision that can be achieved in extracting the cross-section, it has limited use in constraining new physics in a simultaneous global fit of several operators. It is worthwhile to make use of other measurements.

[¶]Note, however, that a full marginalisation will be overly conservative when confronting a concrete UV model.

4.4.2 Polarised beams

One of the principal strengths of lepton colliders is that the polarisation of the incoming beams can be finely controlled, so that the relative contributions between different sub-processes to a given final state can be tuned. Moreover, because the dependence of top observables on the operators of Eq. (4.10) depends strongly on the initial state polarisation, varying the settings increases the number of independent measurements that can be used to place bounds in a global fit.

To emphasise this point, we study the forward-backward asymmetry, defined as

$$A_{FB}^t = \frac{N(\cos\theta_t > 0) - N(\cos\theta_t < 0)}{N(\cos\theta_t > 0) + N(\cos\theta_t < 0)}, \tag{4.12}$$

where θ_t is the polar angle between the top quark and the incoming electron, for three incoming beam polarisation settings: unpolarised beams, denoted $(A_{FB}^t)_U$; a fully left-handed initial polarised electron beam and fully right-handed polarised positron beam, denoted $(A_{FB}^t)_L$; and vice versa, denoted $(A_{FB}^t)_R$. The SM predictions for these settings at tree level are $\{(A_{FB}^t)_U, (A_{FB}^t)_L, (A_{FB}^t)_R\} \simeq \{0.40, 0.37, 0.47\}$, which agree well with the full NNLO QCD estimates [390, 391]. The dependence of these asymmetries on the operators of Eq. (4.10) is shown in Fig. 4.10.

We see that the dependence on the operators distinctively depends on the initial state polarisations. For the $(A_{FB}^t)_L$ case, we again see the large interference-square cancellation in the gauge-type operators $\mathcal{O}_{uW}$ and $\mathcal{O}_{uB}$. For the right-handed case the impact of $\mathcal{O}_{uW}$ is much milder. For both cases we see that the operators $\mathcal{O}_{\varphi q}^{(3)}$ and $\mathcal{O}_{\varphi q}^{(1)}$ pull the prediction in opposite directions. Most encouragingly, we see that the departure from the SM prediction is now much stronger for the φ-type operators than for the total cross-section, which should lead to a sizeable improvement in the final constraints.

To generate these constraints, we consider a global fit of the four operators to six observables:

$$\{(A_{FB}^t)_U, (A_{FB}^t)_L, (A_{FB}^t)_R, (\sigma_{\text{tot}}^{t\bar{t}})_U, (\sigma_{\text{tot}}^{t\bar{t}})_L, (\sigma_{\text{tot}}^{t\bar{t}})_R\}. \tag{4.13}$$

In extracting the constraints, we consider the more realistic ILC polarisation capabilities $\mathcal{P}_{e^-} = \pm 0.8$, $\mathcal{P}_{e^+} = \mp 0.3$, noting that the cross-section for arbitrary e^+e^- polarisations is related to the fully polarised one by [392, 393]

$$\sigma_{\mathcal{P}_{e^-}\mathcal{P}_{e^+}} = \frac{1}{4}\{(1+\mathcal{P}_{e^-})(1-\mathcal{P}_{e^+})\sigma_{\text{RL}} + (1-\mathcal{P}_{e^-})(1+\mathcal{P}_{e^+})\sigma_{\text{LR}}\}, \tag{4.14}$$

where σ_{RL} is the cross-section for fully right-handed polarised electrons and fully left-handed polarised positrons and σ_{RL} is vice versa (the σ_{RR} and σ_{LL} components vanish for p-wave annihilation into spin-1 bosons). Performing a χ^2 fit of the full analytic expression

for each observable, using SM pseudodata with 1% experimental error bars (based on studies in Refs. [383,394]) and SM theory uncertainties of 1% (based on the calculations of Refs. [378,379,388,390,391]) the individual and marginalised constraints on these operators are shown in Fig. 4.11.

At the level of individual operators, the constraints are not improved drastically by adding in asymmetry information. For the global fit, however, the constraints lead to much stronger bounds than for fitting to cross-sections (although the marginalisation typically weakens the overall constraints by a factor $\mathcal{O}(100)$).

We see that the constraints are again much stronger for the field strength operators $\mathcal{O}_{uW}$ and $\mathcal{O}_{uB}$, where the constraints are at the $|\bar{c}_i| \lesssim 10^{-4}$ level for the individual constraints and $|\bar{c}_i| \lesssim 10^{-2}$ for the marginalised case, corresponding to a mass reach of $\Lambda \gtrsim 10$ TeV and $\Lambda \gtrsim 2.16$ TeV, respectively, assuming $c_i \simeq 1$. The weakest constraints are on the operators $\mathcal{O}_{\varphi q}^{(3)}$ ($\mathcal{O}_{\varphi q}^{(1)}$), which translate into bounds on Λ of roughly 700 GeV.

While it is encouraging that the bounds are consistent with an EFT formulation, in the sense that $\Lambda \gg \sqrt{s}$, the ILC mass reach for the scale of new physics that would generate these operators is still low. We note, however, that these bounds are on the conservative side, since other observables such as oblique parameters and LEP asymmetries contribute complementary information that will in general tighten them. To keep this fit self-contained, we postpone this discussion until Sec. 4.5.

4.4.3 CLIC constraints

The Compact Linear Collider (CLIC) project [395,396], with its larger maximum centre-of-mass energy $\sqrt{s} = 3$ TeV, will be in a stronger position to discover the effects of some higher-dimensional operators, whose effects naïvely scale with the CM energy as s/Λ^2. There are two main running scenarios, but both envisage total integrated luminosities of 500 fb^{-1} at $\sqrt{s} = 500$ GeV, 1.5 ab^{-1} at 1.4 or 1.5 TeV, and 2 ab^{-1} at 3 TeV. Again, we focus on the highest energy setting $\sqrt{s} = 3$ TeV, to maximise discovery potential for non-resonant new physics through $D = 6$ operators.

Moving further away from the $t\bar{t}$ threshold, the total $e^+e^- \to Z^*/\gamma \to t\bar{t}$ rate is smaller than at the ILC; at $\sqrt{s} = 3$ TeV it is around 20 fb, which means for the total forecast integrated luminosity at this energy there will be a statistical uncertainty of $\simeq 0.5\%$. A total experimental uncertainty of 1% may therefore be too optimistic an estimate once systematics are fully itemised. Nonetheless, for ease of comparison with the ILC figures, we take this as a baseline, and the corresponding constraints, using the same observables and beam settings, are shown in Fig. 4.12.

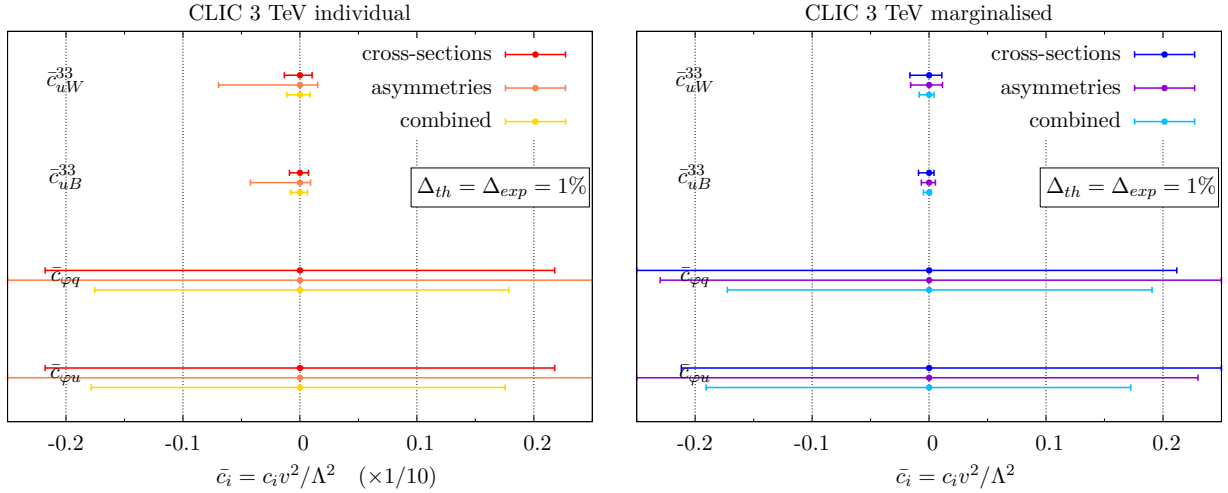We see that for the individual fit, CLIC constraints are of the same order of magnitude

Figure 4.12: *95% confidence ranges for the operators we consider here, from CLIC running at $\sqrt{s} = 3$ TeV, assuming 1% theoretical and experimental uncertainties, by fitting to cross-sections, asymmetries, and the combination, with each operator considered individually (a) or in a 5D fit (b). To display both on the same plot, we scale the individual constraints up by a factor of 10, so that the bottom axis is actually $\bar{c}_i/10$.*

worse than ILC ones.[||] Although the direct sensitivity to the operators is enhanced, we see that as we move away from the $t\bar{t}$ threshold, the interference effect of the φ-type operators is much smaller. This is not the case for the operators $\mathcal{O}_{uW}$ and $\mathcal{O}_{uB}$, whose contributions stem mainly from the $(D = 6)^2$ term, as seen on the right of Fig. 4.9, which receives no suppression. Their individual constraints are close to the ILC values, indicating that energy scale is not the dominant factor driving these limits, but rather the theory and experimental uncertainties which saturate the sensitivity, which we do not vary.

For the more general marginalised fit, we see again that combining cross-section and asymmetry measurements will break blind directions in the fit, leading to much more powerful overall constraints. Unlike for the case of the ILC, however, care must be taken in interpreting these limits in terms of the mass scale of a particular UV model. The marginalised constraint $|\bar{c}_{\varphi u}| \lesssim 0.05$, for example, corresponds to a mass scale $\Lambda/\sqrt{c} \gtrsim 1.1$ TeV, which is less than the energy scale probed in the interaction, so that the constraint can only be linked to a particular model if it is very weakly coupled: $g_* \ll 1$.

[||]This is in contrast to Higgs sector constraints from $e^+e^- \to hZ$, where the projected sensitivity is extremely dependent on the momentum flow through the vertex, leading to better overall CLIC constraints [397].
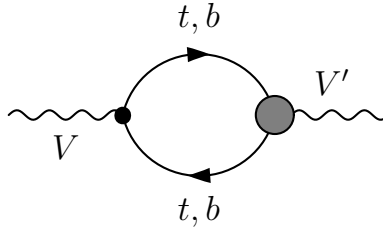
Figure 4.13: *Representative Feynman diagram contributing to the S, T, U parameters at one-loop. The grey-shaded area marks a possible dimension six insertion while the black dot represents a SM vertex of the $V - V'$ polarisation function, $V, V' = W^\pm, Z, \gamma$.*

4.5   Beyond $e^+e^- \to t\bar{t}$: Precision Observables

Obviously the direct constraints that we have focused on in this work do not exist in a vacuum and the interplay of direct and indirect sensitivity plays an important part in ultimately obtaining the best constraints for a given model (see [318, 325]). To put the expected direct constraints detailed above into perspective we analyse the impact of the considered operators on LEP precision observables. Note, that these Z resonance observables are sensitive to a plethora of other new interactions and a direct comparison is not immediately straightforward [325]. Nonetheless, there is significant discriminative power that is worthwhile pointing out, which we will discuss in the following.

4.5.1   Oblique corrections

The S, T, U parameters [398, 399] (see also [400]) are standard observables that capture oblique deviations in the SM electroweak gauge sector from the SM paradigm [401, 402] through modifications of the gauge boson two-point functions. The operators considered in this work modify these at the one-loop level through diagrams of the type shown in Fig. 4.13. Throughout we perform our calculation in dimensional regularisation.

   The definitions of S, T, U, see [398, 399, 402], are such that in the SM all divergencies cancel when replacing the renormalised polarisation functions by their bare counterparts. The modifications of Fig. 4.13, however, induce additional divergencies due to the dimension 6 parts and the introduction of two-point function counterterms is essential to obtain a UV-finite result, see also [318]. This leads to a regularisation scale μ_R dependence of the $D = 6$ amplitude parts after renormalisation, as shown in Fig. 4.14. It is this part which we focus on as we choose the SM with a 125 GeV Higgs as reference point [403].
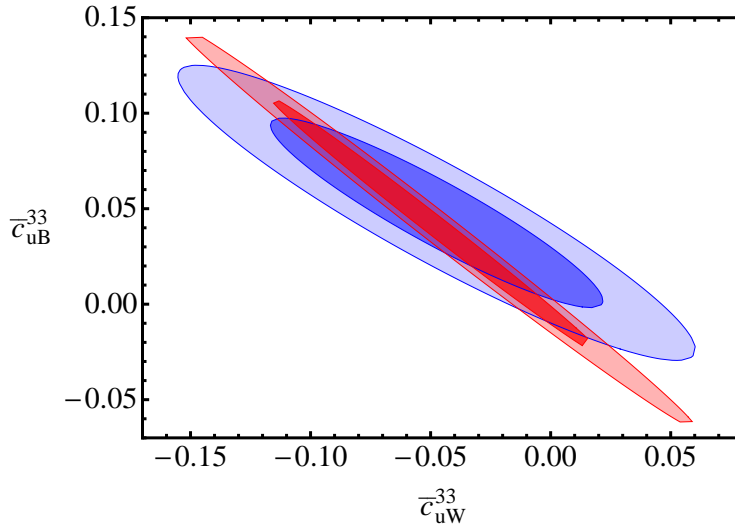
Figure 4.14: *Contour of the S, T, U fit reported in [403] for specifically the operators $\mathcal{O}_{uW}, \mathcal{O}_{uB}$, which are unconstrained by down-sector measurements. All other Wilson coefficients are chosen to be zero. The dark and light shaded areas represent 68% and 95% confidence levels for this projections, while the blue contour uses $\mu_R = m_Z$ and the red contour $\mu_R = 1$ TeV.*

4.5.2 Non-oblique corrections

A well-measured quantity at LEP is the $Zb\bar{b}$ vertex, which enters the prediction of the bottom forward-backward asymmetry $A_{FB}^{b\bar{b}}$, see e.g. [404]. Similar to the operators in Eq. (4.10), in the generic dimension six approach we can expect similar operators for the down-sector of the 3rd fermion family. These will modify the interactions along the same lines as we focused on above for the top sector. However, due to the different isospin properties, the bottom forward backward asymmetry is now sensitive to the sum $c_{\varphi q}^{(3)} + c_{\varphi q}^{(1)}$. This leads to a complementary constraint by the LEP forward backward asymmetry compared to the direct measurements in $t\bar{t}$, as shown on the left of Fig. 4.15.

Moreover, the constraints on $c_{\varphi q}^{(3)} + c_{\varphi q}^{(1)}$ from $A_{FB}^{b\bar{b}}$ can be combined with the constraints on $c_{\varphi q}^{(3)} - c_{\varphi q}^{(1)}$ to extract independent bounds on $c_{\varphi q}^{(3)}$ and $c_{\varphi q}^{(1)}$. This is shown on the right of Fig. 4.15. Care should be taken when interpreting these constraints individually, however. We are considering marginalised bounds for the ILC constraint but only one operator combination for the LEP bound. In general, other operators that we do not consider here will impact the $Zb\bar{b}$ vertex at tree level and in general weaken the bound. This serves as a useful visualisation, however, of the complementarity between past and future colliders in constraining these operators.
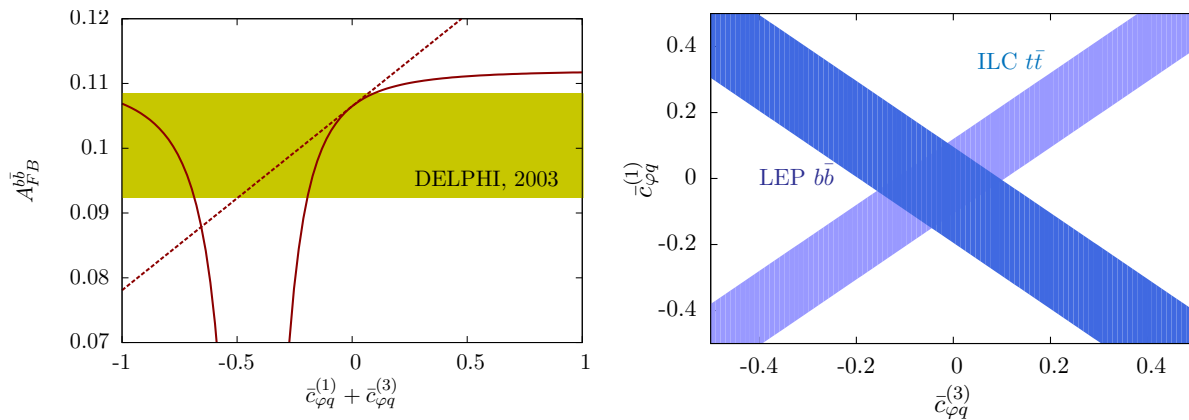
Figure 4.15: *Left: Forward-backward asymmetry (linearised, dashed, and full results, solid) as a function of $c_{\varphi q}^{(3)} + c_{\varphi q}^{(1)}$. The exclusion contour is taken from DELPHI collaboration's Ref. [404] for the most constraining measurement at $\sqrt{s}$ =91.26 GeV. Right: Allowed 95% confidence regions for the Wilson coefficients $\bar{c}_{\varphi u}^{(3)}$ and $\bar{c}_{\varphi u}^{(1)}$ obtained from combining the information from ILC $t\bar{t}$ asymmetries and cross-sections (dark blue) and LEP1 $b\bar{b}$ measurements (lighter blue).*

4.6   Summary

Given the unsatisfactory precision of current probes of top quark electroweak couplings from hadron collider measurements, they must be a key priority in the physics agenda of any future linear e^+e^- collider. By parameterising non-standard top couplings through $D = 6$ operators, we have analysed the potential for the ILC and CLIC to improve the current precision of the top electroweak sector. Unsurprisingly, if experimental precision would match current estimates, and theory uncertainties can be brought to the same level, the current constraints can be drastically improved by both colliders, with associated bounds on the scale of new physics typically in the 1 TeV to few TeV range, depending on the assumed coupling structures of the underlying model. Using asymmetry measurements as well as cross-sections will be crucial to this endeavour, as will collecting large datasets with several incoming beam polarisations.

We have found that, unlike for the Higgs sector, the large increase in centre-of-mass energy at CLIC does not necessarily offer a competitive advantage over the ILC for bounding new top interactions by the operators we consider, and bounds on the operators that we consider are typically stronger at the latter, though in simultaneous 4D fits the difference is not striking. For some of the operators we consider, the bounds derived from CLIC fits correspond to mass scales smaller than the CM energy that we consider, which can call into question the validity of the EFT description, unless the CLIC sensitivity can exceed the expectations we forecast here.

By combining Z-pole measurements from LEP1 with $t\bar{t}$ measurements (and future improved electroweak precision measurements), one can in principle break degeneracies in the operator set and disentangle individual operators that could previously be only bounded in combinations. We showed this for the LEP forward-backward asymmetry, this could be improved by fitting other precision electroweak observables too. Care must be taken in interpreting the associated constraints, however, as both sets of measurements will in general talk to other operators for which there is no complementarity, and a more systematic approach taking into account EFT loop corrections would have be undertaken before these numerical bounds can be taken at face value.

5 Top tagging with Deep Neural Networks

5.1 Introduction

In the last chapter, we discussed how the sensitivity to non-resonant new physics in the top quark sector could be improved by making use of reconstruction techniques for 'boosted' [405–410] top final states; that is, top quarks whose decay products are highly collimated in the detector. Since we generally expect on dimensional grounds that the effects of heavy new physics coupling to the top quark would be most prominent in the high p_T tail, top tagging is well suited to this task. There are many other applications of boosted tagging, however, and as higher momentum transfer final states begin to be probed more regularly at the LHC, jet substructure methods are becoming an increasingly crucial component of its analysis program.

In top pair production for instance, 15% of the (Standard Model) total cross-section comes from the region where $p_T^t \gtrsim 200$ GeV, and in new physics scenarios where, for example, a heavy resonance decays into a top pair, the boosted regime is the region of interest [411]. Because they rely on the final state objects being well-separated in the detector, standard 'resolved' reconstruction techniques begin to falter here. Boosted reconstruction techniques, on the other hand, which allow for more than one of the hard partons in the final state to be captured by the same large-radius jet, and then analyse the substructure of those jet(s), are more efficient in this regime. Improving the efficiency of boosted top tagging in the increasingly challenging hadronic environment of the LHC detectors will thus be of vital importance over the LHC lifetime. This is the subject of this chapter.

There exist several well-studied examples of these boosted techniques, for example mass-drop filtering [406], trimming [412], pruning [413, 414], shower deconstruction [415] and the HEPTopTagger [343, 344, 361] algorithm, which was used in chapter 4. These taggers have typically been designed to focus on the hard substructures of the jet, and veto the softer activity in the detector. This approach, whilst well-motivated from a QCD perspective, in principle throws out valuable information about the jet's properties, and it is interesting to ask whether the softer constituents of the jets can also offer powerful discriminating features between signal and background, as has already been demonstrated for the hard substructures.

An intriguing new avenue in this direction has recently been opened, that makes use of machine learning algorithms known as Convolutional Neural Networks (ConvNets). These *deep learning* techniques are routinely used in computer vision problems such as image/facial recognition as well as natural language processing. As applied to boosted jet finding, the basic idea is to view the calorimeter (η, ϕ) plane as a sparsely filled image, where

the filled pixels correspond to the calorimeter cells with non-zero energy deposits and the pixel intensities to the energy or E_T deposited. After some image preprocessing, a training sample of these *jet images*, with signal and background events, is fed through a ConvNet designed to learn the signal-like and background-like features of the images; shapes, clusters, edges, voids, etc. The final layer of the network converts the learned features of the image into a probability of it being either signal or background. One then feeds a test sample through the network to quantify its performance, usually expressed in terms of its receiver operator characteristic (ROC) curve: signal efficiency vs. background rejection. This method will be described in detail later in the chapter.

Using Monte Carlo simulated data, these networks have been shown to be comparable in performance to various already well-established QCD-inspired taggers, even in the presence of pileup [416–424]. However, these analyses have typically focused on hadronically decaying W-boson tagging, and top tagging has not yet been studied. Whilst a conceptually similar problem, the extra prong in the top decay corresponding to the b-quark, and the extra mass splitting in the tWb vertex, lead to a different structure of the final large-R jet. It would be desirable to show that jet image techniques can be applied here with similar efficiency. A broader goal would be to demonstrate that these techniques have applicability across the realm of jet substructure, not just in the narrow example of W-tagging. This analysis is a step in that direction.

This chapter is structured as follows: In section 5.2 we outline the main jet substructure and top tagging concepts visited throughout the analysis. In section 5.3 we discuss the two machine learning algorithms that we utilise: boosted decision trees (BDTs) and convolutional neural networks. In section 5.4 we discuss the details of how we build a jet image and the specific network architectures we test against. In section 5.5 we discuss the performance of the neural network compared to standard QCD based taggers, and investigate the physics that the network learns, before summarising in section 5.6.

5.2   QCD-inspired top tagging

The goal of any QCD-based tagger is to use an understanding of perturbative QCD to construct observables and algorithms that offer the best discrimination between jets from signal processes; typically involving tops, W and Z bosons or Higgses, and background; typically originating from QCD dijet or multijet processes. The simplest such observable is the *jet mass*, defined simply as the mass of the Lorentz vector sum of all of the constituent momenta in a jet.

$$m_J^2 = \sum_i^{\text{constits}} p_i^\mu p_{\mu i}. \tag{5.1}$$

Clearly, the exact value of the reconstructed jet mass will be dependent on the algorithm used to cluster the jet, but it should be close to the mass of the parton(s) it originated from. In the case of signal jets, this should be fairly close to the mass of the decaying object, as shown on the left of Fig. 5.1, where the reconstructed fat jet mass for jets within the range $350\text{GeV} \leq p_T^J \leq 450\text{GeV}$ is shown (to ensure the jets are tagged as tops we require that they are matched to a parton-level top within $\Delta R = 1.2$). In the case of QCD jets, hard perturbative emissions push the reconstructed jet mass to higher values than one would naïvely expect from its massless constituents, as shown on the right hand side of Fig. 5.1, where the QCD jet mass is broadly peaked at $m_J \sim 100$ GeV, rather than towards the origin. Some additional processing of the reconstructed jets therefore must be done before the jet mass can be considered a reliable observable, i.e. a detector-level observable that captures the kinematics of the hard process. These techniques are generally referred to as *grooming*.

5.2.1 Jet grooming

The three most commonly used jet groomers are *filtering*, *pruning* and *trimming*. Filtering a jet is aimed at keeping its hard constituents originating from the hard interaction whilst rejecting as much as possible the softer constituents originating from QCD radiation. Beginning with a large-R jet, its constituents are reclustered with the Cambridge-Aachen algorithm (C/A) [358, 359], with radius parameter R_{filt}. Then, all constituents outside of the N hardest subjets are filtered out. That is, the N subjets with the largest transverse momenta after filtering are kept, the rest of the jet is discarded. For a hadronically decaying top quark $t \to Wb \to q\bar{q}'b$, there are three hard subjets associated with the decay, and one typically allows tolerance for some extra QCD emission, to avoid mis-filtering out the decay products. The 5 hardest subjets are usually kept. A typical radius R_{filt} is 0.3, for an original jet of size $R = 1.2$ or 1.5. The effects of filtering on the jet mass for top and QCD jets are shown in Fig. 5.1. We see that the mass peak for the signal jets has become much sharper, and is symmetric around the top mass. For the QCD jets, we see a moderate shift towards zero.

For trimming, rather than keeping the N hardest subjets, one cuts on the fraction of the subjet p_T to the original fat jet p_T. Namely, one reclusters the fat jet constituents with the k_T algorithm using a radius R_{sub}, thus clustering softer energy deposits first and harder activity last. Any subjets i with $p_{Ti}/p_T^J < f_{\text{cut}}$ are removed, and the final trimmed jet is composed of the remaining subjets. The jet mass distributions for tops and QCD after trimming, using $R_{\text{sub}} = 0.35$, $f_{\text{cut}} = 0.03$, are also shown in Fig. 5.1. Comparing to the raw distributions, we see that QCD jets typically lose 30%-50% of their mass, whereas top jets keep most of their mass, therefore trimming is a robust procedure for isolating the hard components of a jet that one is most interested in. A drawback is that sometimes the
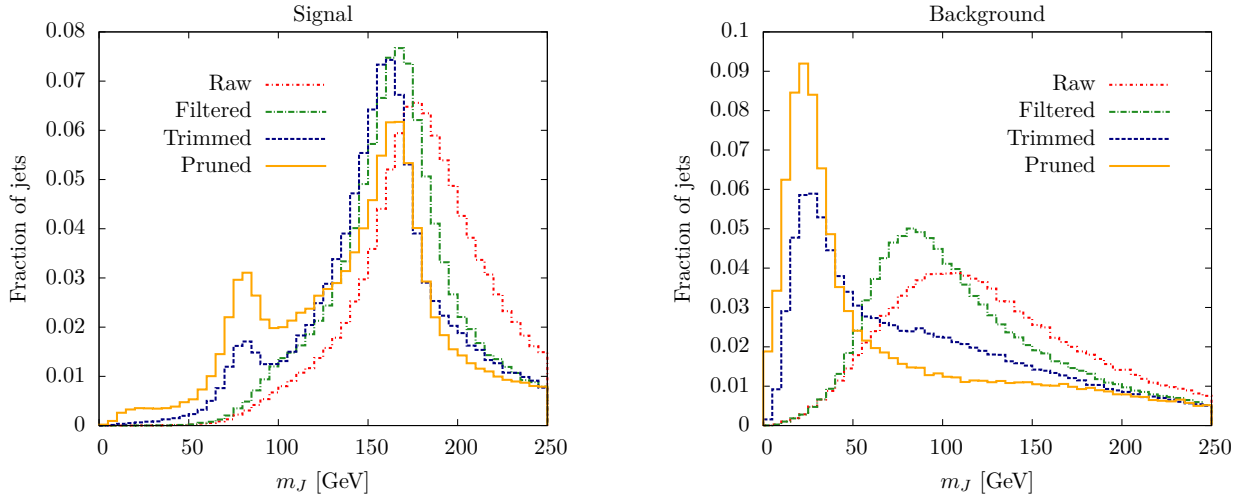
Figure 5.1: *Top (left) and QCD (right) jet mass distributions after the application of various grooming algorithms. The original jets are constructed with the Cambridge-Aachen algorithm with distance parameter 1.5. Filtering is applied with radius parameter $R_{filt} = 0.3$ and the 5 hardest subjets are kept. The trimming procedure uses $R_{sub} = 0.35$ and $f_{cut} = 0.03$, and the pruned jets are reclustered with the C/A algorithm with $z_{cut} = 0.1$, $R_{cut} = 0.5$ (see text for details).*

b quark will be vetoed from the jet if it is very soft, and so the jet mass will be spuriously reconstructed to the W mass.

Pruning is related to trimming, but along with kinematic requirements on the p_T, it applies an additional geometrical cut, designed to remove wide-angle radiation. Starting with a large-R jet, one reapplies either the C/A or k_T [425, 426] jet clustering algorithm on its constituents, but at each clustering step, the criterion that a) the softer of the constituents j_1 and j_2 (taken to be j_2, so that $p_T^{j_1} > p_T^{j_2}$) has $p_T^{j_2}/p_T^{j_1+j_2} < z_{\text{cut}}$, and b) $\Delta R_{j_1,j_2} < R_{\text{cut}} \times (2m_J/p_T^J)$, are both tested.

At least one of a) or b) must be true, otherwise j_1 and j_2 are both removed. Typical values used in real jet substructure analyses are $z_{\text{cut}} = \{0.05, 0.1\}$ and $R_{\text{cut}} = \{0.1, 0.2, 0.3\}$. The effects of pruning on the signal and background jets we are using are also shown in Fig. 5.1. We see that pruning is particularly effective for reducing the QCD jet mass, but the price paid is that it is also more likely to erroneously reconstruct the signal jet mass to the W mass instead of the top, i.e. it is more likely to veto the b quark.

5.2.2 N-subjettiness

As well as the absolute mass scales entering the jet, it is also useful to construct observables that capture the likely number of hard subjets within a jet, i.e. that gives a measure of
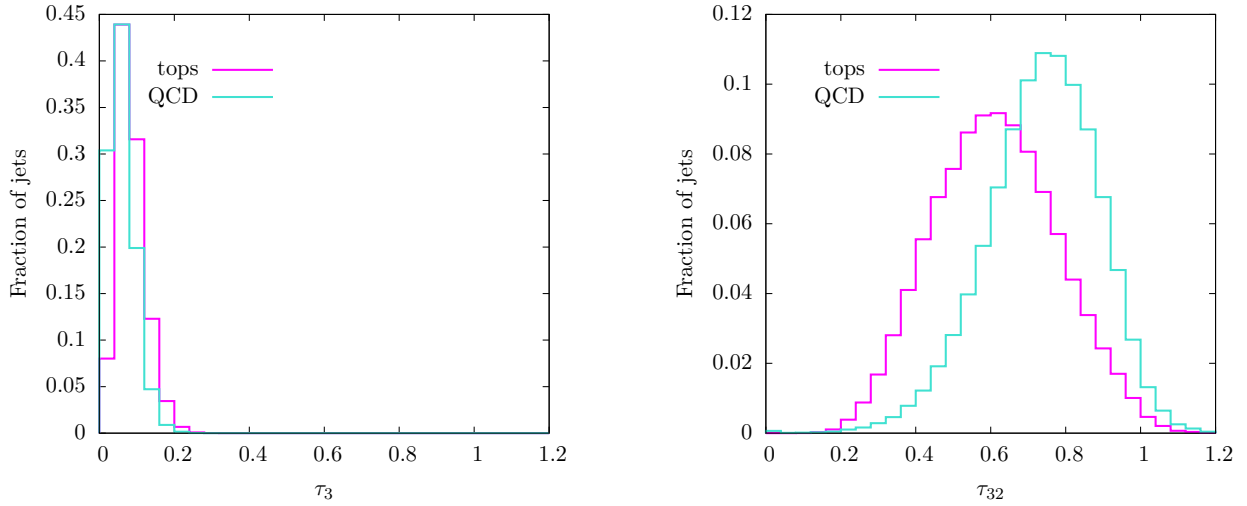
Figure 5.2: *Distributions of the N-subjettiness parameter τ_3 (left) and ratio τ_3/τ_2 (right) for top and QCD jets, clustered using the C/A algorithm with $R = 1.5$ as before. The subjet axes are defined by k_T clustering, and $\beta = 1$. Better discriminating power is offered by the ratio τ_3/τ_2, as seen on the right plot.*

how 'prongy' a fat jet is. One could imagine, for instance, a W' or Z' boson with a mass close to the top quark. It would have a characteristic two-prong decay, but from jet mass distributions alone it would appear very similar to a top jet. N-subjettiness [427, 428] is an observable designed to measure this property. For a fat jet of radius R_0, N-subjettiness is defined as

$$\tau_N = \frac{\sum_i p_{Ti} \min(\Delta R_{i1}, \ldots, \Delta R_{iN})^\beta}{\sum_i p_{Ti} R_0}. \tag{5.2}$$

The sum is over the jet constituents i, and N is the number of candidate subjets. To define the distance measures ΔR, one must define candidate subjet axes. This can be done either by summing over all possible candidate subjet directions and taking the minimum, or, in a computationally simpler approach, by defining the subjets by applying the k_T clustering algorithm on the jet constituents and truncating after N subjets are generated. The exponent β is a free parameter that can be optimised for a given analysis. It is also implicitly assumed that the fat jet radius $R_0 > \Delta R_{ij}$ for all jet constituents.

In the limit $\tau_N \to 0$, the jet must have $\Delta R_{iN} = 0$ for all constituents, i.e. all constituents must be perfectly aligned with the candidate subjets, so the jet has exactly N subjets. In the case of $\tau_N \to 1$, the jet must have at least $N + 1$ subjets, i.e. the minimisation missed some subjet axes. Therefore jets with small τ_N are said to be N-subjetty, whereas jets with larger τ_N have more than N subjets. For top jets the ratio τ_3 is clearly the one of interest. In fact, it has been shown that the ratio $\tau_3/\tau_2 \equiv \tau_{32}$ has the most discriminating power between tops and QCD, because several QCD uncertainties are present in the value

of τ_3 that drop out when taking the ratio τ_{32}*. The distributions of τ_{32} for tops and QCD dijets are shown in Fig. 5.2, showing clear separation between signal and background. N-subjettiness is thus an extremely useful discriminating variable and will be utilised throughout this analysis.

5.2.3 SoftDrop

Another useful substructure algorithm for top tagging is SOFTDROP [429], which is designed to iteratively remove soft, wide-angle activity corresponding to contamination from the underlying event from the jet. The algorithm is as follows:

- For a fat jet J, undo the last step of the clustering so that there are two subjets j_1 and j_2.

- If the softer of the two subjets has a fraction of the total jet p_T greater than

$$\frac{\min(p_T^1, p_T^2)}{p_T^1 + p_T^2} > z \left(\frac{\Delta R_{12}}{R_0} \right)^\beta,$$

(5.3)

  then J is the final jet.

- If not, keep harder of the two jets and continue until the final jet is reached.

The parameters of SOFTDROP are thus the exponent β and the p_T fraction z. The limit $\beta \to 0$ removes the geometrical dependence of the cut, e.g. $\beta = 0$, $z = 0.1$ removes subjets with less than 10% of the total jet p_T. The additional ΔR cut shows that softer activity is much more likely to be cut out if it has a large angular separation from the rest of the jet. SOFTDROP is a useful procedure for defining a jet mass that is robust against contamination from soft radiation, and a 'softdropped' jet mass, in conjunction with τ_{32} provides powerful discrimination between tops and QCD background.

There are many other observables used in top tagging, and we will not review them all here, we merely briefly discuss the ones that will be utilised in this analysis. The other key component of the analysis is the use of machine learning techniques. In the form of multivariate techniques such as Boosted Decision Trees, these techniques have been well-established as useful in jet substructure classification. Deep Neural Networks are a recent development. Since we will benchmark the performance of our neural network against various standard BDTs, in the next section we will outline the basics of a BDT architecture, before discussing the structure of a deep neural network in detail.

---

*See Ref. [345] for a discussion of this, and as a useful review of top tagging in general.

5.3 Machine learning inspired top tagging

5.3.1 Boosted decision tree approaches

Starting with a sample of signal and background jets that one wants to classify, a standard decision tree begins by ordering the input jets by the value of some discriminator variables, typically referred to as *features* in BDT parlance. For our case, the features will be quantities such as the jet mass and N-subjettiness ratios. For each feature, the sample is split into two parts, based on the value of the feature that best separates the signal and background. An initial node, containing all the events, has thus been separated into two branches. The branches continue to be split until the final branches, called *leaves* are either pure signal or pure background, or contain too few samples of either kind to continue splitting.

Each sample jet/event is given a weight W_i, where the weighting procedure is specific to the precise type of decision tree classifier. The splitting criterion is defined by the purity of the sample in each branch,

$$P = \frac{\sum_s W_s}{\sum_s W_s + \sum_b W_b}, \tag{5.4}$$

For each branch, the optimal splitting of the parent branch into two daughter branches is defined by maximising the *Gini* function

$$[(\sum_{i=1}^{n} W_i)P(1-P)]_{\text{parent}} - [(\sum_{i=1}^{n} W_i)P(1-P)]_{\text{daughter 1}} - [(\sum_{i=1}^{n} W_i)P(1-P)]_{\text{daughter 2}}. \tag{5.5}$$

If the events all have unit weight, then if a final leaf has a purity $P \geq 1/2$ it is called a signal leaf, and if $P \leq 1/2$ it is a background leaf[†]. All events that finish on a signal leaf are classified as signal, likewise for background. The signal efficiency is defined as the number of signal events that landed on signal leaves divided by the initial number of signal samples, and the corresponding background contamination or 'mistag' rate is the number of background events that landed on signal leaves divided by the initial number of background samples. A schematic diagram of a simple two-feature decision tree is shown in Fig. 5.3.

The performance of a decision tree classifier can be significantly enhanced by applying 'boosting' criteria at each branch. A boosted decision tree proceeds as above, except if a signal event is misclassified as background or vice versa, the weight of that event is *boosted*. The entire sample of weighted events is fed through a second tree, and the procedure

---

[†]In fact the purity demanded can be varied between zero and one, which gives a smooth curve of background contamination for a given signal efficiency, i.e. an ROC curve.
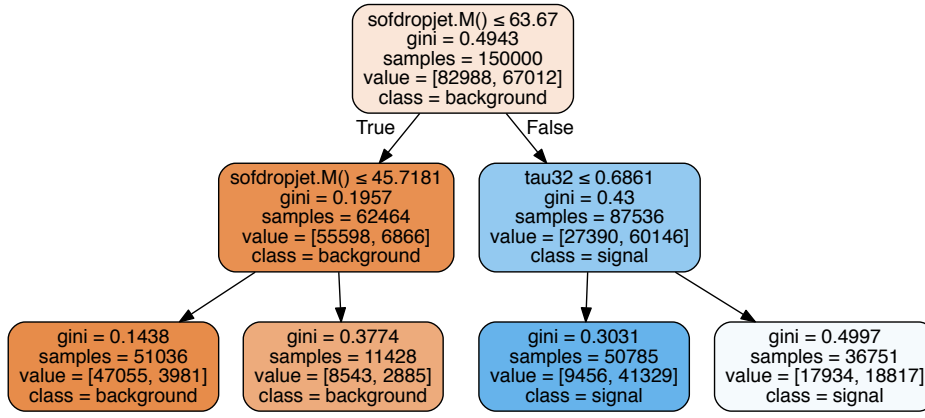
Figure 5.3: *A simple decision tree classifier with a maximum tree depth of two, using just two features: the N-subjettiness ratio τ_{32} and the mass of the fat jet after applying the* SOFTDROP *algorithm. The tree first cuts on the jet mass, with jets above the cut going into the signal category, a further cut on τ_{32} is then applied to the signal jets, and a further cut on the jet mass to the background jets. Figure generated with [430].*

repeated for several trees. After each tree, the event is given a score of 1 if it lands on a signal leaf and -1 if it lands on a background leaf. The total BDT score for that event is the sum of its scores after each tree.

A commonly used boosting classifier is the adaptive boosting method or AdaBoost. Starting with a sample of N events labeled i, each with a weight $1/N$, give signal events an initial score $y_i = 1$ and background events $y_i = -1$. For a set of input features x_i for each event i, define $T_m(x_i) = 1$ if the event lands on a signal leaf at the end of tree m and $T_m(x_i) = -1$ if it lands on a background leaf. A 'misclassify' score I is defined as $I(y_i \neq T_m(x_i)) = 1$ and $I(y_i = T_m(x_i)) = 0$, so that the total error on the mth tree is just the weighted sum of the misclassified event scores:

$$err_m = \frac{\sum_i^N w_i I(y_i \neq T_m(x_i))}{\sum_i^N w_i}. \tag{5.6}$$

The misclassified events are reweighted by

$$w_i \rightarrow w_i \times \exp\left(\alpha_m I(y_i \neq T_m(x_i))\right), \tag{5.7}$$

where

$$\alpha_m = \beta \times \log\left(\frac{1 - err_m}{err_m}\right). \tag{5.8}$$

The parameter β can be chosen freely; typical values used are $\beta = 1$ or $\beta = 1/2$. The goal of the decision tree is then to simply construct the weights at each tree so that

the error function is minimised for each event. Boosted decision trees are typically more powerful because their performance is less sensitive to changing the model parameters, and they are less prone to overfitting. For this reason they are already widely used in signal/background discrimination in real particle physics analyses, and will thus provide an appropriate comparison for quantifying the performance of a deep neural network.

5.3.2 Deep learning

Deep neural network algorithms originate from computer vision, namely the problem of trying to correctly label an image from a fixed set of categories. The network consists of three steps: input, learning and evaluation.

An input image can be thought of as a vector of pixel densities $\{x_i\}$. The goal is then to construct a classifier which converts this into a label from a predefined set $\{y_i\}$. In a binary classification problem such as signal vs. background, $\{y_i\} \in \{0, 1\}$. The simplest such classifier is a linear mapping, where each pixel is multiplied by a series of weights W and added with some bias b: $y_i = W_{ij}x_j + b_i$. For a binary classifier over a (flattened) $N \times N$ pixel image, W is then a $2 \times N^2$ matrix and b a 2D vector. This transformation would constitute a single *layer* of a neural network. A typical NN has many such layers, with in general more complicated mappings between layers.

Each layer in a network is made of of *units* or *neurons* which receive as input the output from the previous layer. If every neuron in the layer is connected to every neuron in the previous layer, the layer is *fully-connected* or *dense*. The connections between neurons in neighbouring layers may be thought of as synapses, in analogy with neurons in the brain. In each neuron, the outputs from the previous layer are combined with the weights and biases of the current layer and compared to an *activation function*, which models whether the signal entering the neuron is strong enough for the neuron to 'fire', i.e. to activate the next layer. After passing the image through several such *hidden layers*, the final layer of the network is a classifier which converts the inputs to a probability that the original image $\vec{x}$ belonged to class y_i. The goal of a neural network is to construct appropriate weights at each layer such that the distinctions between the different classes are maximally 'learned' by the neurons.

In order to construct the weights and biases, one trains the network by feeding through a set of input images for which the classes are known. A *loss function* quantifies how close the network class prediction for the image is to the truth label on the data. The entire network can thus be thought of as a single, differentiable, highly nonlinear function acting on the input pixel densities to produce an output classification. There are many choices of network available. 'Deep learning' is distinguished from ordinary neural networks by large numbers of layers, but each layer is typically connected to a few neighbouring neurons in

the previous layer, rather than shallow networks with small numbers of dense layers. As it is believed to simulate the process of image recognition in the human brain, it is well suited to the problem of image classification. Convolutional Neural Networks are among the most widely used deep learning architectures. Here we discuss the main building blocks of a ConvNet and their effects and discuss the training of the network.

The convolutional neural network starts from a two-dimensional input image and identifies characteristic patterns using a stack of convolutional layers. We use a set of standard operations, starting from the $n \times n$ image input I:

- ZeroPadding: $(n \times n) \to (n + 2 \times n + 2)$
  We artificially increase the image by adding zeros at all boundaries in order to remove dependence on non-trivial boundary conditions,

$$I \to \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & I & \vdots \\ 0 & \cdots & 0 \end{pmatrix} . \tag{5.9}$$

- Convolution: $n'_{\text{c-kernel}} \times (n \times n) \to n_{\text{c-kernel}} \times ((n - n_{\text{c-size}} + 1) \times (n - n_{\text{c-size}} + 1))$
  To identify features in an $n \times n$ image or feature map we linearly convolute the input with $n_{\text{c-kernel}}$ kernels of size $n_{\text{c-size}} \times n_{\text{c-size}}$. If in the previous step there are $n'_{\text{c-kernel}} > 1$ layers, the kernels are moved over all input layers. For each kernel this defines a feature map $\widetilde{F}^k$ which mixes information from all input layers

$$\widetilde{F}^k_{ij} = \sum_{l=0}^{n'_{\text{c-kernel}}-1} \sum_{r,s=0}^{n_{\text{c-size}}-1} \widetilde{W}^{kl}_{rs} \, I^l_{i+r,j+s} + b_k \qquad \text{for} \quad k = 0, ..., n_{\text{c-kernel}} - 1 . \tag{5.10}$$

- Activation: $(n \times n) \to (n \times n)$
  This non-linear element allows us to create more complex features. A common choice is the rectified linear activation function (ReL) which sets pixel with negative values to zero, $f_{\text{act}}(x) = \max(0, x)$. In this case we define for example

$$F^k_{ij} = f_{\text{act}}(\widetilde{F}^k_{ij}) = \max\left(0, \widetilde{F}^k_{ij}\right) . \tag{5.11}$$

Instead of introducing an additional unit performing the activation, it can also be considered as part of the previous layer.

- Pooling: $(n \times n) \to (n/p \times n/p)$
  We can reduce the size of the feature map by dividing the input into patches of fixed

size $p \times p$ (sub-sampling) and assign a single value to each patch

$$F'_{ij} = f_{\text{pool}}\left(F_{(ip\ldots(i+1)p-1,\,jp\ldots(j+1)p-1)}\right) . \tag{5.12}$$

MaxPooling returns the maximum value of the subsample $f_{\text{pool}}(F) = \max_{\text{patch}}(F_{ij})$.

A convolutional layer consists of a ZeroPadding, Convolution, and Activation step each. We then combine $n_{\text{c-layer}}$ of these layers, followed by a pooling step, into a block. Each of our $n_{\text{c-block}}$ blocks therefore works with essentially the same size of the feature maps, while the pooling step between the blocks strongly reduces the size of the feature maps. This ConvNet setup efficiently identifies structures in two-dimensional jet images, encoded in a set of kernels W transforming the original picture into a feature map. In a second step of our analysis the ConvNet output constitutes the input of a fully connected DNN, which translates the feature map into an output label y:

- Flattening: $(n \times n) \to (n^2 \times 1)$
  While the ConvNet uses two-dimensional inputs and produces a set of corresponding feature maps, the actual classification is done by a DNN in one dimension. The transition between the formats reads

$$x = (F_{11}, \ldots, F_{1n}, \ldots, F_{n1}, \ldots, F_{nn}) . \tag{5.13}$$

- Fully connected (dense) layers: $n^2 \to n_{\text{d-node}}$
  The output of a standard DNN is the weighted sum of all inputs, including a bias, passed through an activation function. Using rectified linear activation it reads

$$y_i = \max\left(0, \sum_{j=0}^{n^2-1} W_{ij} x_j + b_i\right) . \tag{5.14}$$

For the last layer we apply a specific SoftMax activation function

$$y_i = \frac{\exp\left(W_{ij} x_j + b_i\right)}{\sum_i \exp\left(W_{ij} x_j + b_i\right)} . \tag{5.15}$$

It ensures $y_i \in [0, 1]$, so the label can be interpreted as a signal or background probability.

In a third step we define a cost or loss function, which we use to train our network to a training data set. For a fixed architecture a parameter point θ is given by the ConvNet weights $\widetilde{W}_{rs}^{kl}$ defined in Eq.(5.10) combined with the DNN weights W_{ij} and biases b_i defined

in Eq.(5.14). The performance of the training is quantified by minimising the mean squared error

$$L(\theta) = \frac{1}{N} \sum_{i=0}^{N} (y(\theta; x_i) - y_i)^2 \ , \tag{5.16}$$

where $y(\theta; x_i)$ is the predicted binary label of the input x_i and y_i is its true value. For a given parameter point θ we compute the gradient of the loss function $L(\theta)$ and first shift the parameter point from θ_n to θ_{n+1} in the direction of the gradient $\nabla L(\theta_n)$. In addition, we can include the direction of the previous change such that the combined shift in parameter space is

$$\theta_{n+1} = \theta_n - \eta_L \nabla L(\theta_n) + \alpha(\theta_n - \theta_{n-1}) \ . \tag{5.17}$$

The learning rate η_L determines the step size and can be chosen to decay with each step (decay rate). The parameter α, referred to as momentum, dampens the effect of rapidly changing gradients and improves convergence. The Nesterov algorithm changes the point of evaluation of the gradient to

$$\theta_{n+1} = \theta_n - \eta_L \nabla L(\theta_n + \alpha(\theta_n - \theta_{n-1})) + \alpha(\theta_n - \theta_{n-1}) \ . \tag{5.18}$$

Each training step (epoch) uses the full set of training events.

5.4 Analysis setup

5.4.1 Building a jet image

The analysis objects for the ConvNet are jet images, constructed from Monte Carlo simulations. For signal events we use 14 TeV LHC $t\bar{t}$ samples in the all hadronic decay channel [‡]. For background we consider a QCD dijet sample, which constitutes the dominant background to this signature. All samples are simulated with PYTHIA8 [431], without the effects of multiparton interactions. The events are then passed through a fast detector simulation with DELPHES3 [432], using the ATLAS card, and calorimeter towers of size $\Delta\eta \times \Delta\phi = 0.1 \times 5°$. We cluster these towers with the anti-k_T [356] algorithm with $R = 1.5$, as implemented in FASTJET3 [357], requiring that all jets have $|\eta| < 1.0$. These anti-k_T jets give us a smooth outer shape of the fat jet and a well-defined jet area for our jet image.

To ensure that the jet substructure in the jet image is consistent with QCD, and to

---

[‡]One could also consider the semileptonic $t\bar{t}$ channel, which has a much cleaner signature, but since we are interested in the properties of the fat jets rather than the event as a whole, we make use of the much higher statistics of the all-hadronic channel.
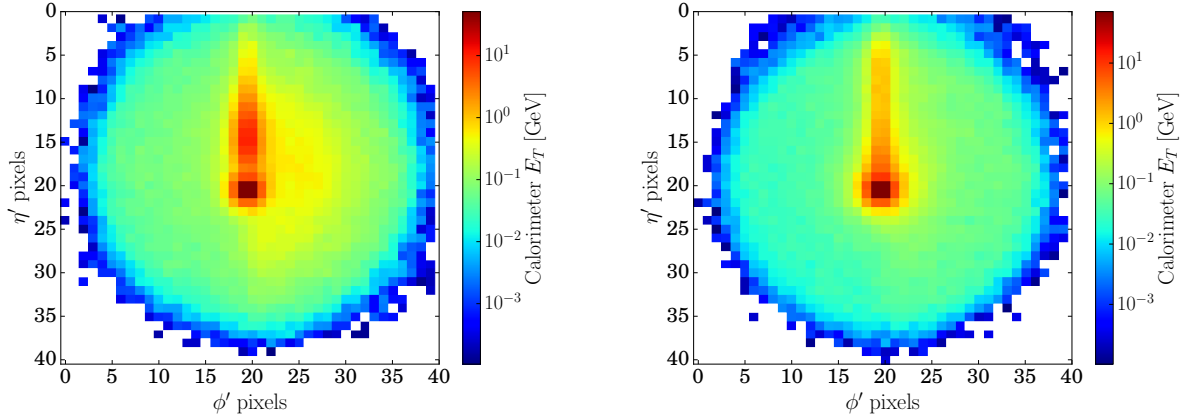
Figure 5.4: *Jet image after pre-processing for the signal (left) and background (right). Each picture is averaged over 10,000 actual images.*

prepare the jets for the HEPTOPTAGGER algorithm, we re-cluster constituents of the anti-k_T jet using the Cambridge-Aachen (C/A) algorithm with $R = 1.5$. Its substructures define the actual jet image pixels. A final comment is in order here, when we identify these calorimeter towers with pixels, it is not clear whether the information used should be the energy E or only its transverse component E_T. We study the performance of the network in both cases.

The rather stringent geometrical cut $|\eta_{\text{fat}}| < 1.0$ guarantees that the fat jets are contained entirely in the central part of the detector and justifies our calorimeter tower size. For this study we focus on the range $p_{T,\text{fat}} = 350 \ldots 450$ GeV, such that all top decay products can be easily captured in the fat jet. For signal events, we additionally require that the fat jet can be associated with a Monte-Carlo truth top quark within $\Delta R < 1.2$.

Before feeding in the images to the neural network, it is helpful to apply some preprocessing, such that the salient features of the image are contained in the same location§. The preprocessing steps are as follows:

1. Find maxima: before we can align any image we have to identify characteristic points. Using a filter of size 3×3 pixels, we localize the three leading maxima in the image.

2. Shift: we then shift the image to center the global maximum taking into account the periodicity in the azimuthal angle direction.

3. Rotation: next, we rotate the image such that the second maximum is in the 12 o'clock position. The interpolation is done linearly.

---

§In facial recognition, this would be akin to shifting the image such that the eyes are always at the centre.
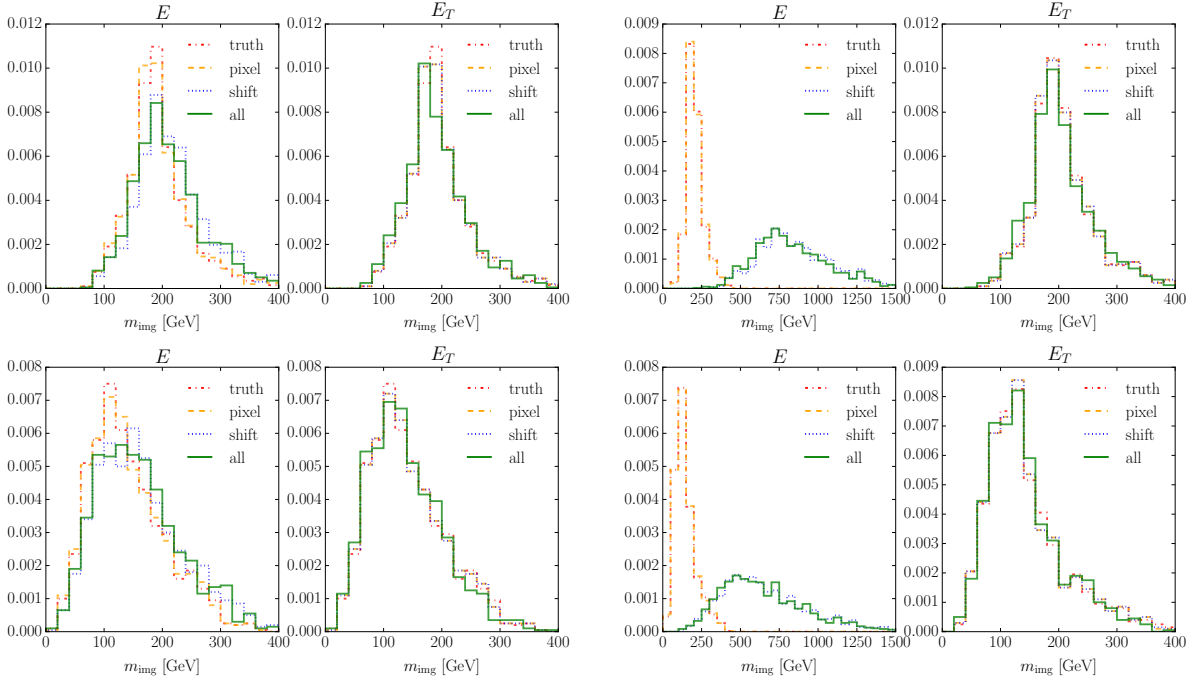
Figure 5.5: *Effect of the preprocessing on the image mass calculated from E-(left) and E_T-images (right) of signal (top) and background(bottom). The right set of plots illustrates the situation for forward jets with $|\eta| > 2$.*

4. Flip: next we flip the image to ensure the third maximum is in the right half-plane.

5. Crop: finally, we crop the image to 40×40 pixels.

Throughout the analysis we will apply two pre-processing setups: for minimal pre-processing we apply steps 1, 2 and 5 to define a centered jet image of given size. Alternatively, for full pre-processing we apply all five steps. In Fig. 5.4 we show averaged signal and background images based on the transverse energy from 10,000 individual images after full pre-processing. The leading subjet is in the center of the image, the second subjet is in the 12 o'clock position, and a third subjet from the top decay is smeared over the right half of the signal images. These images indicate that fully pre-processed images might lose a small amount of information at the end of the 12 o'clock axis.

As mentioned above, a non-trivial question is whether one should use the calorimeter tower E or E_T as the pixel density, since the shift and rotation steps of pre-processing involve a longitudinal η boost, under which E is not invariant. Following Ref. [416] we investigate the effect on the mass information contained in the images,

$$m_{\text{img}}^2 = \left[\sum_i E_i \left(1, \ \frac{\cos \phi_i'}{\cosh \eta_i'}, \ \frac{\sin \phi_i'}{\cosh \eta_i'}, \ \frac{\sinh \eta_i'}{\cosh \eta_i'} \right) \right]^2 \qquad E_i = E_{T,i} \cosh \eta_i', \qquad (5.19)$$
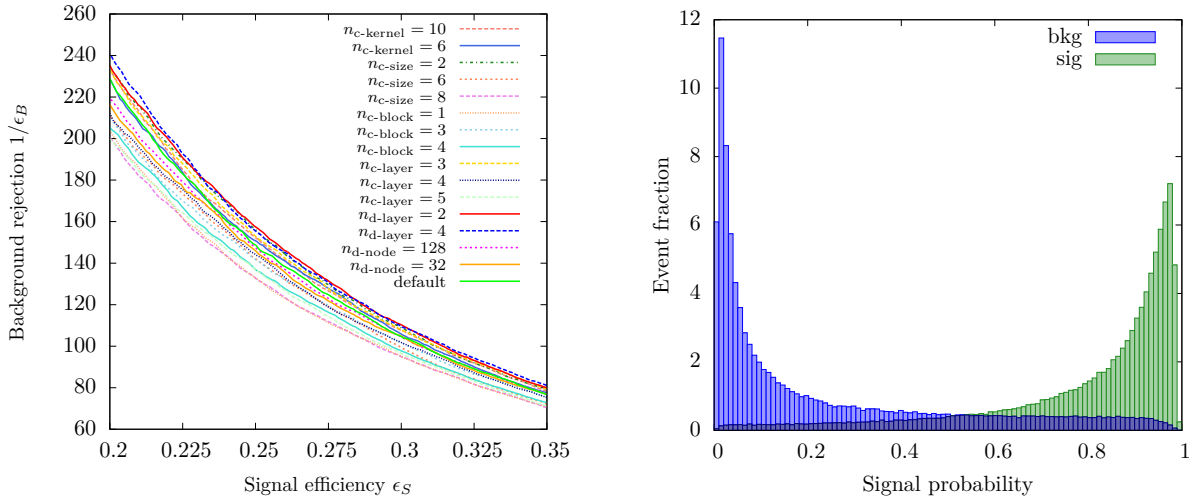
Figure 5.6: *Left: performance of some of our tested architectures for full pre-processing in terms of an ROC curve, including the default* DEEPTOP *network. Right: discrimination power or predicted signal probability for signal events and background probability for background events. We use the default network.*

where η_i' and ϕ_i' are the centre of the ith pixel after pre-processing. The study of all pre-processing steps and their effect on the image mass in Fig. 5.5 illustrates that indeed the rapidity shift has the largest effect on the E images, but this effect is not large. For the E_T images the jet mass distribution is unaffected by the shift pre-processing step. The reason why our effect on the E images is much milder than the one observed in Ref. [416] is our condition $|\eta_{\text{fat}}| < 1$. In the the right-hand panels of Fig. 5.5 we illustrate the effect of pre-processing on fat jets with $|\eta| > 2$, where the image masses changes dramatically. Independent of these details we use pre-processed E_T images as our machine learning input [433]. Since, as a rule of thumb, neural networks perform better with small numbers, we scale the images such that the pixel entries are between 0 and 1.

5.4.2 Network architecture

The neural network architecture is constructed with KERAS [434] using the THEANO backend [435]. To identify a suitable DEEPTOP network architecture, we scan over several possible realizations or hyper-parameters. As discussed in the last section, we start with jet images of size 40×40. For architecture testing we split our total signal and background samples of 600,000 images each into three sub-samples. After independently optimizing the architecture we train the network with 150,000 events and after each training epoch test it on an independent test sample of the same size. The relative performance on the training and test samples allows us to avoid over-training. Finally, we determine the performance of the default network on a third sample, now with 300,000 events.

In a first step we need to optimize our network architecture. The ConvNet side is organized in $n_{\text{c-block}}$ blocks, each containing $n_{\text{c-layer}}$ sequences of ZeroPadding, Convolution and Activation steps. For activation we choose the ReL step function. Inside each block the size of the feature maps can be slightly reduced due to boundary effects. For each convolution we globally set a filter size or convolutional size $n_{\text{c-size}} \times n_{\text{c-size}}$. The global number of kernels of corresponding feature maps is given by $n_{\text{c-kernel}}$. Two blocks are separated by a pooling step, in our case using MaxPooling, which significantly reduces the size of the feature maps. For a quadratic pool size of $p \times p$ fitting into the $n \times n$ size of each feature map, the initial size of the new block's input feature maps is $n/p \times n/p$. The final output feature maps are used as input to a DNN with $n_{\text{d-layer}}$ fully connected layers and $n_{\text{d-node}}$ nodes per layer.

In the left panel of Fig. 5.6 we show the performance of some test architectures. We give the complete list of tested hyper-parameters in Tab. 5.1. As our default we choose one of the best-performing networks after explicitly ensuring its stability with respect to changing its hyper-parameters. The hyper-parameters of the default network we use for fully as well as minimally pre-processed images are given in Tab. 5.1. In Fig. 5.7 we illustrate this default architecture.

In the second step we train each network architecture using the mean squared error as our loss function and the Nesterov algorithm with an initial learning rate $\eta_L = 0.003$. We train our default setup over up to 1000 epochs and use the network configuration minimizing the loss function calculated on the test sample. Different learning parameters were used to ensure convergence when training on the minimally pre-processed and the scale-smeared samples. Because the DNN output is a signal and background probability, the minimum signal probability required for signal classification is a parameter that allows to link the signal efficiency ϵ_S with the mis-tagging rate of background events ϵ_B.

In Sec. 5.5 we will use this trained network to test the performance in terms of ROC

hyper-parameter	scan range	default
$n_{\text{c-block}}$	1,2,3,4	2
$n_{\text{c-layer}}$	2,3,4,5	2
$n_{\text{c-kernel}}$	6,8,10	8
$n_{\text{c-size}}$	2,4,6,8	4
$n_{\text{d-layer}}$	2,3,4	3
$n_{\text{d-nodes}}$	32,64,128	64
p	0,2,4	2

Table 5.1: *Range of parameters defining the combined ConvNet and DNN architecture, leading to the range of efficiencies shown in the left panel of Fig. 5.6 for fully pre-processed images.*
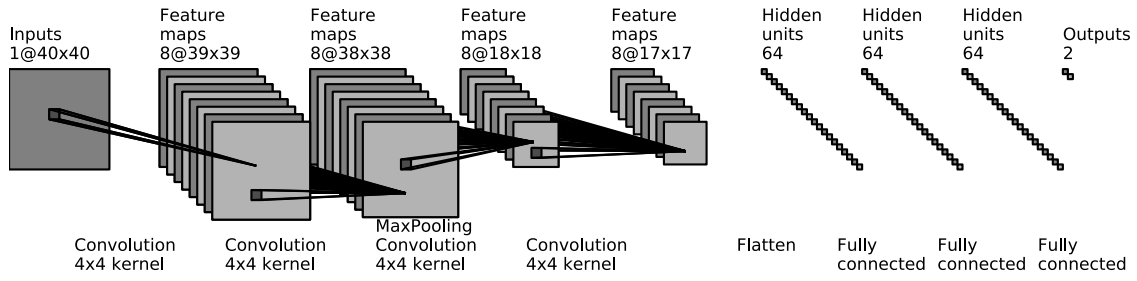
Figure 5.7: *Architecture of our default networks for fully pre-processed images, defined in Tab. 5.1. Figure drawn using Ref. [436].*
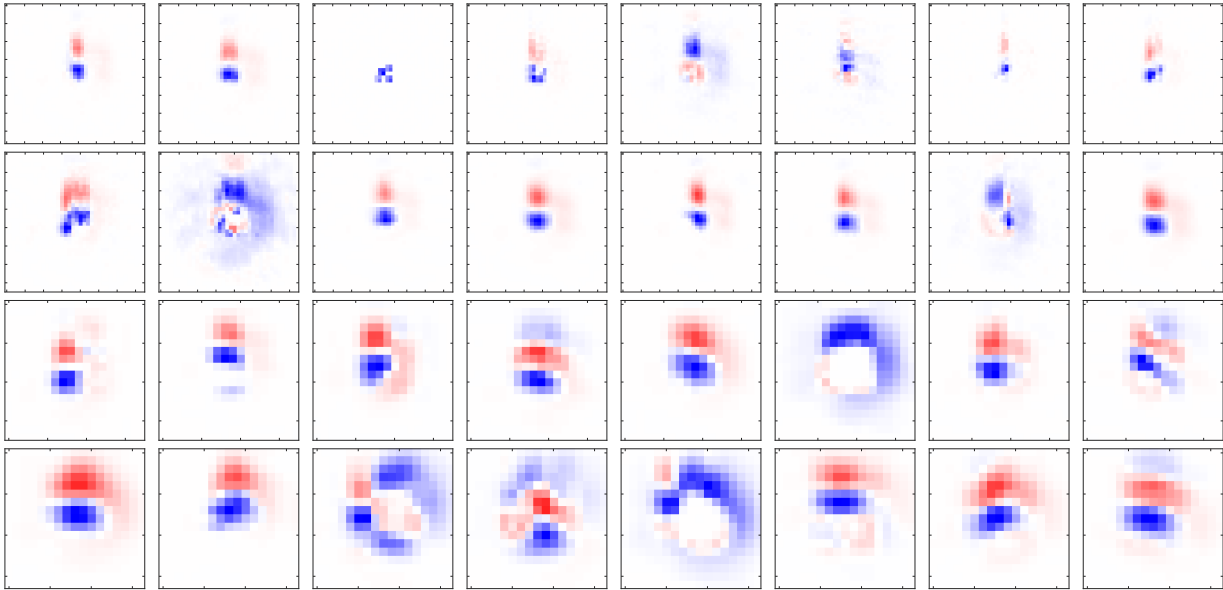
Figure 5.8: *Averaged signal minus background for our default network and full pre-processing. The rows correspond to ConvNet layers one to four. After two rows Max-Pooling reduces the number of pixels by roughly a factor of four. The columns indicate the feature maps one to eight. Red areas indicate signal-like regions, blue areas indicate background-like regions.*

curves, correlating the signal efficiency and the mis-tagging rate.

Before we move to the performance study, we can get a feeling for what is happening inside the trained ConvNet by looking at the output of the different layers in the case of fully pre-processed images. In Fig. 5.8 we show the difference of the averaged output for 100 signal and 100 background images. For each of those two categories, we require a classifier output of at least 0.8. Each row illustrates the output of a convolutional layer. Signal-like red areas are typical for jet images originating from top decays; blue areas are typical for backgrounds. The first layer seems to consistently capture a well-separated second subjet, and some kernels of the later layers seem to capture the third signal subjet in the right half-plane. However, one should keep in mind that there is no one-to-one correspondence
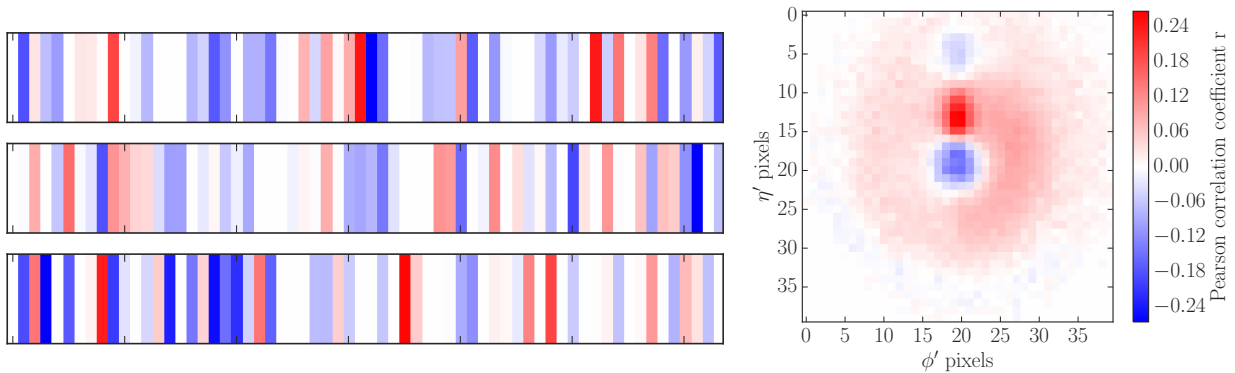
Figure 5.9: *Left: Averaged signal minus background for our default network and full pre-processing. The rows show the three dense DNN layers. Red areas indicate signal-like regions, blue areas indicate background-like regions. Right: Pearson correlation coefficient for 10,000 signal and background images each. The corresponding jet image is illustrated in Fig. 5.4. Red areas indicate signal-like regions, blue areas indicate background-like regions.*

between the location in feature maps of later layers and the pixels in the input image.

On the left hand side of Fig. 5.9 we show the same kind of intermediate result for the two fully connected DNN layers. Each of the 64 linear bars represents a node of the layer. We see that individual nodes are quite distinctive for signal and background images. The fact that some nodes are not discriminative indicates that in the interest of speed the number of nodes could be reduced slightly. The output of the DNN is essentially the same as the probabilities shown in the right panel of Fig. 5.6, ignoring the central probability range between 20% and 80%.

To see which pixels of the fully pre-processed 40×40 jet image have an impact on the signal vs background label, we can correlate the deviation of a pixel x_{ij} from its mean value $\bar{x}_{ij}$ with the deviation of the label y from its mean value $\bar{y}$. A properly normalized correlation function for a given set of combined signal and background images can be defined as

$$r_{ij} = \frac{\sum_{\text{images}} (x_{ij} - \bar{x}_{ij}) (y - \bar{y})}{\sqrt{\sum_{\text{images}} (x_{ij} - \bar{x}_{ij})^2} \sqrt{\sum_{\text{images}} (y - \bar{y})^2}} \ . \tag{5.20}$$

It is usually referred to as the Pearson correlation coefficient. From the definition we see that for a signal probability y positive values of r_{ij} indicate signal-like patterns. On the right hand side of Fig. 5.9 we show this correlation for our network architecture. A large energy deposition in the centre leads to classification as background. A secondary energy deposition in the 12 o'clock position combined with additional energy deposits in the right half-plane lead to a classification as signal. This is consistent with our expectations after full pre-processing, shown in Fig. 5.4.
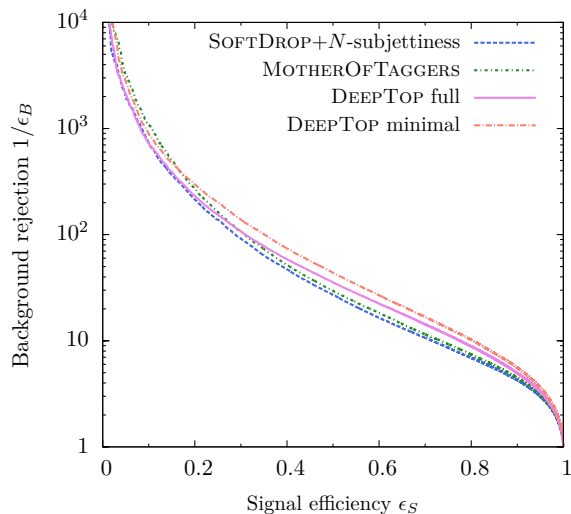
Figure 5.10: *Performance of the neural network tagger compared to the QCD-based approaches* SoftDrop *plus N-subjettiness and including the* HEPTopTagger *variables.*

5.5 Results and benchmarks

Given our optimized machine learning setup introduced in Sec. 5.3 and the fact that we can understand its workings and trust its outcome, we can now compare its performance with state-of-the-art top taggers. The details of the signal and background samples and jet images are discussed in Sec. 5.4; essentially, we attempt to separate a top decay inside a fat jet from a QCD fat jet including fast detector simulation and for the transverse momentum range $p_{T,\text{fat}} = 350 \dots 450$ GeV. Other transverse momentum ranges for the fat jet can be targeted using the same DNN method.

Because we focus on a comparing the performance of the DNN approach with the performance of standard multivariate top taggers we take our Monte Carlo training and testing sample as a replacement of actual data. This means that for our performance test we do not have to include uncertainties in our Pythia simulations compared to other Monte Carlo simulations and data, see Ref. [419] for a study in this direction.

5.5.1 Performance of the network

To benchmark the performance of our DeepTop DNN, we compare its ROC curve with standard Boosted Decision Trees based on the C/A jets using SoftDrop combined with N-subjettiness. From Fig. 5.6 we know the spread of performance for the different network architectures for fully pre-processed images. In Fig. 5.10 we see that minimal pre-processing actually leads to slightly better results, because the combination or rotation and cropping described in Sec. 5.4 leads to a small loss of information. Altogether, the band of different
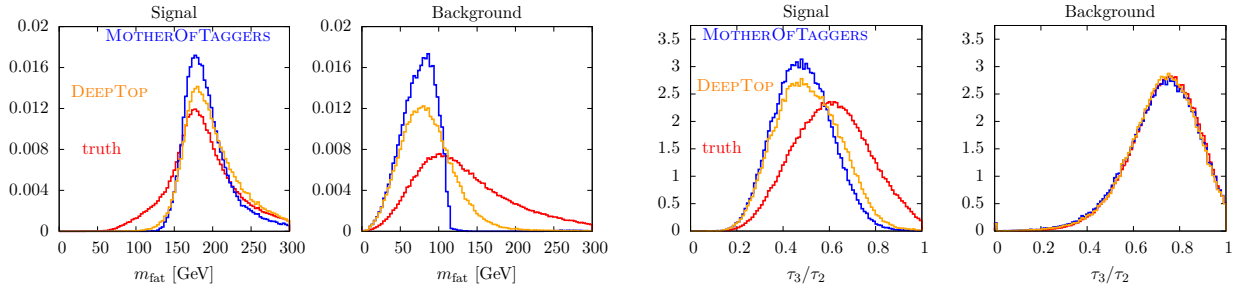
Figure 5.11: *Kinematic observables m_{fat} and τ_3/τ_2 for events correctly determined to be signal or background by the* DEEPTOP *neutral network and by the* MOTHEROFTAGGERS *BDT, as well as Monte Carlo truth.*

machine learning results indicates how large the spread of performance will be whenever for example binning issues in $p_{T,\text{fat}}$ are taken into account, in which case we we would no longer be using the perfect network for each fat jet.

For our BDT we use GRADIENTBOOST in the Python package SKLEARN [430] with 200 trees, a maximum depth of 2, a learning rate of 0.1, and a sub-sampling fraction of 90% for the kinematic variables

$$\{ \, m_{\text{sd}}, m_{\text{fat}}, \tau_2, \tau_3, \tau_2^{\text{sd}}, \tau_3^{\text{sd}} \, \} \qquad (\text{SOFTDROP} + N\text{-subjettiness}) \,, \qquad (5.21)$$

where m_{fat} is the un-groomed mass of the fat jet. This is similar to standard experimental approaches for our transverse momentum range $p_{T,\text{fat}} = 350 \ldots 400$ GeV. In addition, we include the HEPTOPTAGGER2 information from filtering combined with a mass drop criterion,

$$\{ \, m_{\text{sd}}, m_{\text{fat}}, m_{\text{rec}}, f_{\text{rec}}, \Delta R_{\text{opt}}, \tau_2, \tau_3, \tau_2^{\text{sd}}, \tau_3^{\text{sd}} \, \} \qquad (\text{MOTHEROFTAGGERS}) \,. \tag{5.22}$$

In Fig. 5.10 we compare these two QCD-based approaches with our best neural networks. Firstly, we see that both QCD-based BDT analyses and the two neural network setups are close in performance. Indeed, adding HEPTOPTAGGER information slightly improves the SOFTDROP+N-subjettiness setup, reflecting the fact that our transverse momentum range is close to the low-boost scenario where one should rely on the better-performing HEPTOPTAGGER. Second, we see that the difference between the two pre-processing scenarios is in the same range as the difference between the different approaches. Running the DEEPTOP framework over signal samples with a 2-prong W' decay to two jets with $m_{W'} = m_t$ and over signal samples with a shifted value of m_t we have confirmed that the neural network setup learns both, the number of decay subjets and the mass scale.

5.5.2 What the network learns

Following up on on the observation that the neural network and the QCD-based taggers show similar performance in tagging a boosted top decay inside a fat jet, we can check what kind of information is used in this distinction. Both for the DNN and for the MOTHEROF-TAGGERS BDT output we can study signal-like learned patterns in actual signal events by cutting on the output label y corresponding to the 30% most signal-like events shown on the right of Fig. 5.6. Similarly, we can select the 30% most background-like events to test if the background patterns are learned correctly. In addition, we can compare the kinematic distributions in both cases to the Monte Carlo truth. In Fig. 5.11 we show the distributions for m_{fat} and τ_3/τ_2, both part the set of observables defined in Eq. (5.22).

We see that the DNN and BDT tagger indeed learn essentially the same structures. The fact that the signal-like features of those distributions are more pronounced than the Monte Carlo truth is linked to our stiff cut on y, which for the DNN and BDT tagger cases removes events where the signal kinematic features are less pronounced. The MOTHEROFTAGGERS curves for the signal are more peaked than the DEEPTOP curves is due to the fact that the observables are exactly the basis choice of the BDT, while for the neutral network they are derived quantities.

For our performance comparison of the QCD-based tagger approach and the neutral network it is crucial that we understand what the DEEPTOP network learns in terms of physics variables. The relevant jet substructure observables differentiating between QCD jets and top jets are those which which we evaluate in the MOTHEROFTAGGERS BDT, Eq. (5.22). To quantify which signal features the DNN and the BDT tagger have correctly extracted we show observables for signal event correctly identified as such, i.e. requiring events with a classifier response y corresponding to the 30% most signal like events. As we can see from Fig. 5.6 this cut value captures a large fraction of correctly identified events. The same we also do for the 30% most background like events identified by each classifier.

The upper two rows in Fig. 5.12 show the different mass variables describing the fat jet. We see that the DNN and the BDT tagger results are consistent, with a slightly better performance of the BDT tagger for clear signal events. For the background the two approaches deliver exactly the same performance. The deviation from the true mass for the HEPTOPTAGGER background performance is explained by the fact that many events with no valid top candidate return $m_{\text{rec}} = 0$. Aside from generally comforting results we observe a peculiarity: the SOFTDROP mass identifies the correct top mass in fewer that half of the correctly identified signal events, while the fat jet mass m_{fat} does correctly reproduce the top mass. The reason why the SOFTDROP mass is nevertheless an excellent tool to identify top decays is that its background distribution peaks at very low values, around $m_{\text{sd}} \approx 20$ GeV. Even for $m_{\text{sd}} \approx m_W$ the hypothesis test between top signal and
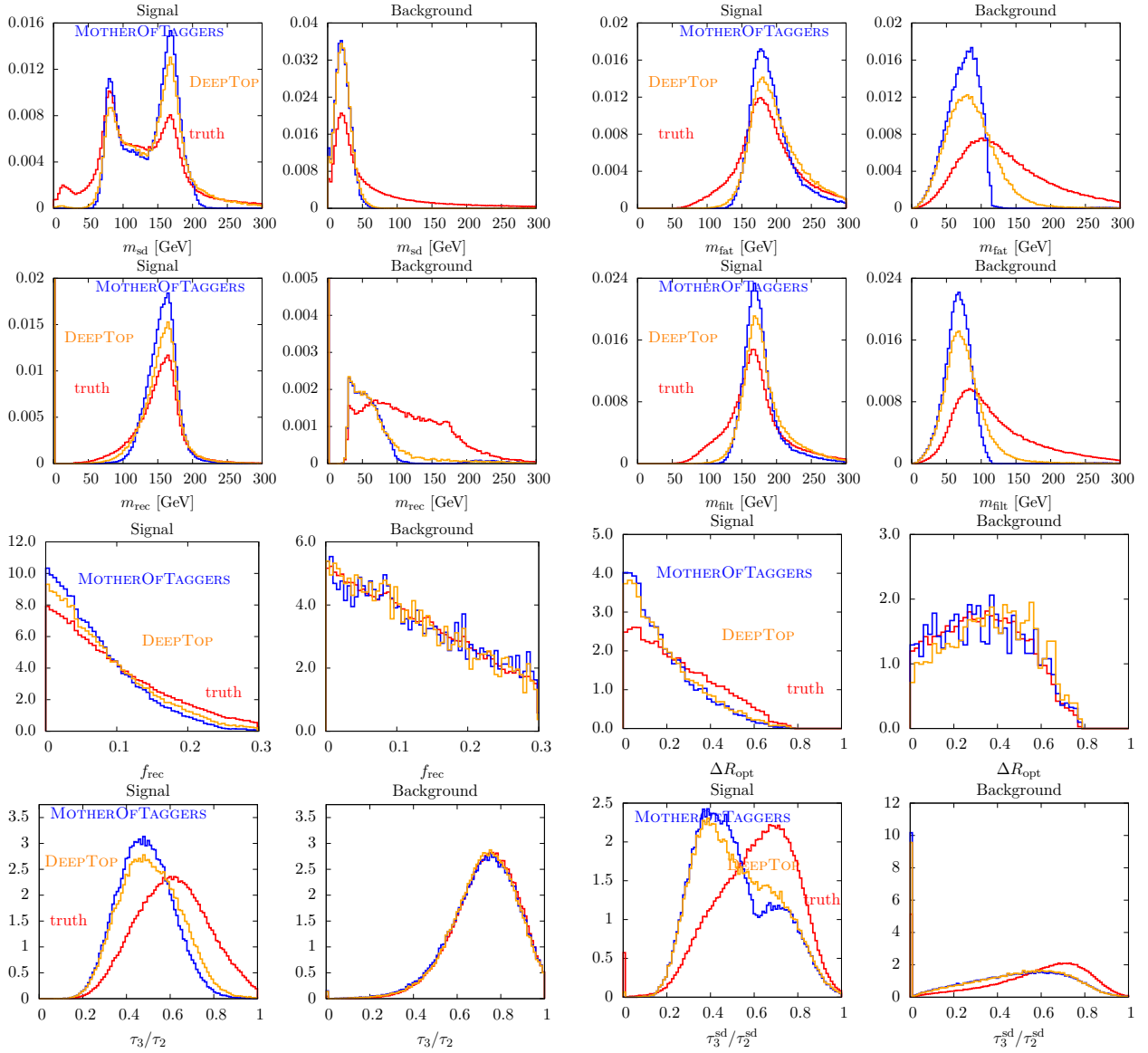
Figure 5.12: *Kinematic observables defined in Eq.(5.22) for events correctly determined to be signal or background by the DeepTop neutral network and by the MotherOfTaggers BDT, as well as Monte Carlo truth. Extended version of Fig. 5.11.*

QCD background can clearly identify a massive particle decay.

In the third row we see that the HepTopTagger W-to-top mass ratio f_{rec} only has little significance for the transverse momentum range studied. For the optimalR variable ΔR_{opt} [361] the DNN and the BDT tagger again give consistent results. Finally, for the N-subjettiness ratio τ_3/τ_2 before and after applying the SoftDrop criterion the results are again consistent for the two tagging approaches.

Following up on the observation that SoftDrop shows excellent performance as a hypothesis test, we show in Fig. 5.13 the reconstructed transverse momenta of the fat jet,
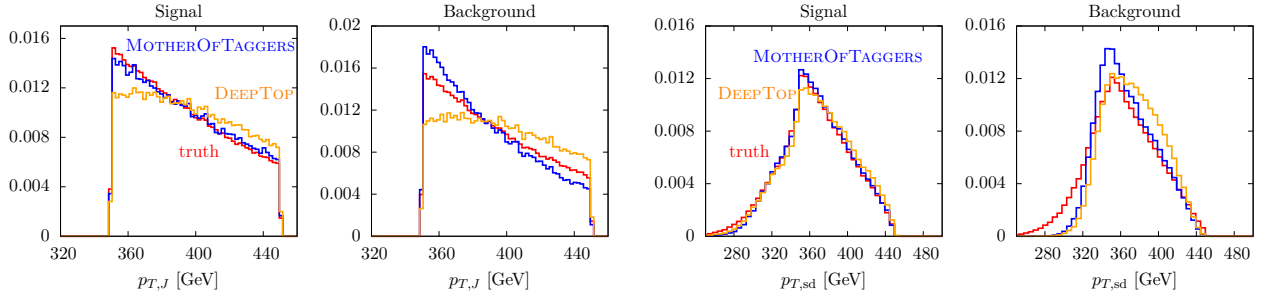
Figure 5.13: *Reconstructed transverse momenta for events correctly determined to be signal or background by the* DEEPTOP *neutral network and by the* MOTHEROFTAGGERS *BDT, as well as Monte Carlo truth.*

or the top quark for signal events. In the left panel we see that the transverse momentum of the un-groomed fat jet reproduces our Monte-Carlo range $p_{T,\text{fat}} = 350 \ldots 450$ GeV. Because the transverse momentum distributions for the signal and background are very similar, the BDT tagger and DNN curves agree very well with the correct behaviour. In the right panel we see that the constituents identified by the SOFTDROP criterion have a significantly altered transverse momentum spectrum. To measure the transverse momentum of the top quark we therefore need to rely on a top identification with SOFTDROP, but a top reconstruction based on the (groomed) fat jet properties.

5.5.3   Sensitivity to experimental effects

Finally, a relevant question is to what degree the information used by the neural network is dominated by low-p_T effects. We can apply a cutoff, for example including only pixels with a transverse energy deposition $E_T > 5$ GeV. This is the typical energy scale where the DNN performance starts to degrade. A key question for the tagging performance is the dependence on the activation threshold. Figure. 5.14 shows the impact of different thresholds on the pixel activation, i.e. E_T used both for training and testing the networks. Removing very soft activity, below 3 GeV, only slightly degrades the network's performance. Above 3 GeV the threshold leads to an approximately linear decrease in background rejection with increasing threshold.

A second, important experimental systematic uncertainty when working with calorimeter images is the calorimeter energy scale (CES). We assess the stability of our network by evaluating the performance on jet images where the E_T pixels are globally rescaled by $\pm 25\%$. As shown in the right panel of Fig. 5.14 this leads to a decline in the tagging performance of approximately 10% when reducing the CES and 5% when increasing the CES.

Next, we train a *hardened* version of the network. It uses the same architecture as
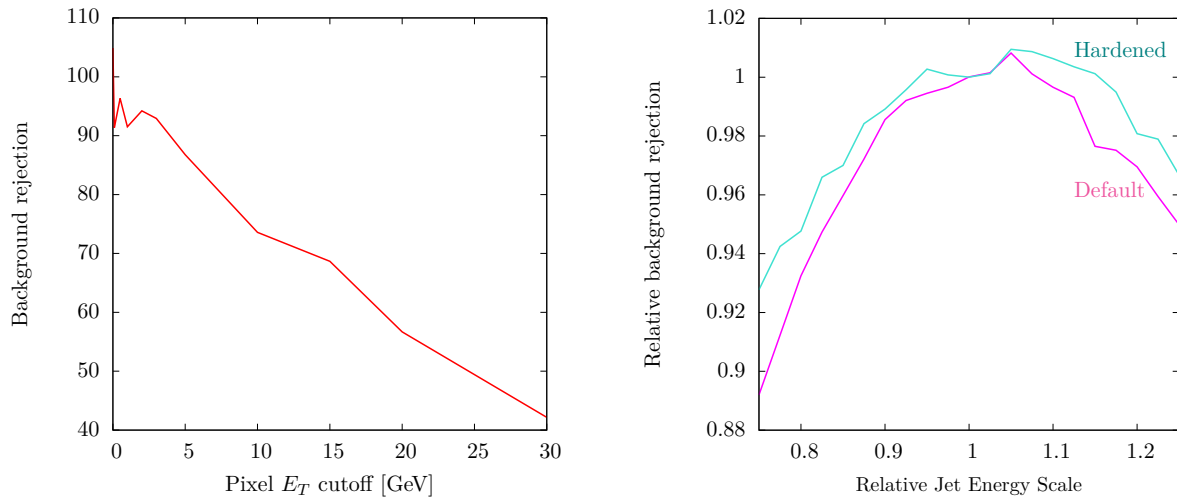
Figure 5.14: *Left: Background rejection at a signal efficiency of 30 % for different activation thresholds. Right: Background rejection for the default and hardened training for different re-scalings of the jet images. The background rejection is evaluated at a signal efficiency of 30 % and normalized to the rejection at nominal calorimeter energy scale.*

our default, but during the training procedure each image is randomly rescaled using a Gaussian distribution with a mean of 1.0 and a width of 0.1. New random numbers are used from epoch to epoch. The resulting network has a similar performance as the default and exhibits a further reduced sensitivity to changes in the global CES. While other distortions of the image, such as non-uniform rescaling, will need to be considered, the resilience of the network and our ability to further harden it are very encouraging for experimental usage where the mitigation and understanding of systematic uncertainties is critical.

5.6 Summary

Like boosted decision trees before them, deep neural networks provide a powerful way of distinguishing between jets from heavy decaying objects, and jets from QCD background. This has already been shown for W-boson tagging, in this analysis we have shown that the same conclusions hold for top tagging as well. Using techniques from image processing, we trained a ConvNet on a Monte Carlo simulated sample of hadronically decaying tops and QCD dijets, and quantified its performance by testing the network on an independent testing sample. By benchmarking the network against the performance of other industry standard taggers, in terms of ROC curves, we showed that the network can offer comparable, and even superior background rejection to QCD inspired taggers. We also showed that preprocessing of the images is an optional, but not necessary step, and good network

performance is observed even on minimally preprocessed images, albeit at the price of more sophisticated network architectures, and thus more intensive computational demands.

Interestingly, we showed that for some observables the network learns features that are not also captured by the BDTs, and vice versa. This shows that the network may be picking up on features of the jets that are not captured by currently used taggers and observables. We have also demonstrated that the network performance is fairly robust against degradation by detector level effects such as calorimeter deposit cutoffs and jet energy scale calibration, which is an encouraging sign that they may one day prove feasible in LHC experimental analyses, as well as phenomenological studies.

Still, there are some shortcomings. Our analysis, like others in this area, is based exclusively on Monte Carlo simulated data samples, for which the training labels are known for each jet. If these techniques are to actually prove useful for the experiments, it will have to be demonstrated that these networks can be trained on real hadron collider data, which, as well as bringing the extra complications of pileup, multiparton interactions and underlying event, does not come with 'truth' labels on each event which can be used to evaluate the network performance. It thus remains to be seen whether or not these techniques will bear fruit in the long term. However, the impressive performance that we see here shows that these techniques certainly merit further study.

Summary and conclusions

This thesis has explored several aspects of top quark phenomenology at hadron and lepton colliders. In chapter 1, we reviewed the foundations of the Standard Model of particle physics, discussed some generalities about hadron collider physics and the main uncertainties that can hinder high theoretical precision there, and explored the properties of the top quark can that can be probed at hadron colliders through its various production mechanisms and decay observables, discussing in detail the state-of-the-art SM theory calculations where relevant.

Chapter 2 focused on physics 'Beyond the Standard Model'. After outlining some of the main flaws of the Standard Model and arguments for new physics (perhaps at the TeV scale); the hierarchy problem, gauge coupling unification and vacuum instability, we presented a few examples of popular new physics models constructed to address these flaws, and we briefly touched upon why they might be relevant for top quark phenomenology. We then moved from specific UV models to a more agnostic approach, by considering the Standard Model as the leading part of an effective theory, where heavy degrees of freedom have been integrated out, leaving behind a tower of higher-dimensional (that is, $D > 4$) operators; which can be studied perturbatively as an expansion in Wilson coefficient divided by UV cutoff $(c_i/\Lambda)^n$.

By using a 'bottom-up' approach, constructing the operators that are consistent with the SM gauge and global symmetries, we derived the only $D = 5$ operator allowed by symmetry constraints. The terms that could generate effects at collider energy scales begin at $D = 6$. Since there are many more operators allowed at this order, we did not reproduce the full derivation of the operator set, we just sketched some of its salient points, following the derivation of the 'Warsaw basis'. Left with 64 non-redundant $D = 6$ operators, we focused on those operators which are (at leading order in the EFT expansion) relevant for top quark physics, which amounted to calculating the effects of the operators on the observables introduced in the first chapter, and computing their numerical effects as a guide for how strongly they may be bounded with current measurements.

In chapter 3, we focused on confronting those operators with data. With the large production rates for top quark associated processes at the LHC and Tevatron, huge statistical samples of top quark data are now public, which make the precision scrutiny of the top quark sector of the Standard Model EFT a timely exercise. The large number of observables requires sophisticated fit machinery, which was achieved by adapting the PRO-FESSOR software package, originally used for tuning of Monte Carlo generators, to BSM limit-setting code. We took various top quark production channels in turn, beginning with $t\bar{t}$ production. We demonstrated that there is complementarity between LHC and Tevatron measurements, because they are dominated by different partonic subprocesses and so most

sensitive to different operators. Differential measurements bring much more sensitivity than total rates alone, as they provide extra sensitivity to shape modifications by the new Lorentz structures of the $D = 6$ operators that go beyond overall normalisation differences. Higher-order processes such as $t\bar{t}Z$ production bring sensitivity to operators that cannot otherwise be constrained, but the statistical uncertainties on the early measurements of these processes are weak and so the subsequent bounds are inconclusive. There is overlap in the operator set constrainable from single top production and from decay observables such as helicity fractions, and so combining these measurements gives stronger bounds.

Some measurements remain in tension with the SM-only prediction, but there are no anomalies that are more significant than what would be naïvely expected with this number of independent measurements. We discussed in detail some of the validity issues inherent in the EFT formulation, such as the neglecting of operators of higher dimension than 6, and the potential for the fit to be excessively pulled by 'overflow' bins in the differential distributions, for which there is no control over scales entering the fit. We concluded that the latter was not a problem in practice by performing the fit with and without those bins, and that for the former, in order to avoid unphysical effects such as negative cross-sections, it is best to keep in contributions from $\mathcal{O}(1/\Lambda^4)$ terms, even when $D = 8$ operators are neglected. Though the fit is comprehensive, and the complementarity between different sets of measurements used allowed the system of operators considered to be easily overconstrained, we found that when our constraints mapped on to the parameter space of specific UV models, the final numerical bounds on the operators considered were in the end rather weak.

Chapter 4 was concerned with the prospects for improving this situation. We noted that the constraints on the $t\bar{t}$ operators, which constituted the largest component of the fit, were dominated by observables reconstructed using 'resolved' techniques in the low to medium p_T region. While statistical uncertainties are smaller in this region, so is the sensitivity to the operators, whose interference typically scales as $\hat{s}/\Lambda^2$. High p_T final states reconstructed by jet substructure methods thus have the potential to dramatically improve the constraints, especially as low statistics in the tails becomes less relevant over the LHC lifetime.

Our analysis showed mixed results, however. While the constraints from a typical low p_T analysis can be improved by up to 70% when current systematics are improved upon and as we approach 3 ab^{-1} of data, the improvement from the boosted region is much milder when experimental error bars are reduced, indicating that SM (and SMEFT) theory uncertainties also have to be dramatically improved in order to exploit the full potential of boosted tagging performance. Given the timescales involved over the forecasted LHC lifetime, however, this does not seem an unrealistic expectation.

We also studied the role that proposed future lepton colliders could play in these im-

provements. By far the weakest constraints from our fit, and the least well-measured of the top quark couplings in general, were from the electroweak neutral vertices. The only direct handle on these from the LHC is from $t\bar{t}Z$ and $t\bar{t}\gamma$ production, and the improvements at 3 ab^{-1} are still modest. Lepton colliders are sensitive to the same operators, however, through the electroweak process $e^+e^- \to t\bar{t}$. We found that orders of magnitude improvement over the current bounds was possible, even when marginalising over all operators in a global fit, which was unsurprising giving the improvement in precision on both the theory prediction and measurement sides.

By benchmarking proposed scenarios at the 500 GeV ILC and 3 TeV CLIC machines, we saw that just as for the LHC, using as much information as possible (in this case running with high statistics at several incoming beam polarisations and making use of pseudo-observables such as forward-backward asymmetries in addition to total cross-sections) was crucial in maximising the sensitivity. Unlike for the LHC, however, we saw that the increasing the collider CM energy does not necessarily buy extra sensitivity, and as we move away from threshold and overall rates become smaller, sensitivity to the operators begins to degrade slightly, although in a full 4D fit the difference was not remarkable. We also highlighted the possibility of combining these bounds with measurements from LEP in order to lift a blind direction in the EFT parameter space, however, at this level of precision, the much more involved EFT loop corrections would have to also be considered in some detail for the numerical bounds to be completely trusted.

In chapter 5, we returned again to hadron colliders, and discussed in detail what goes into the algorithms behind boosted taggers such as the HEPTOPTAGGER, which we used in the boosted analysis of chapter 4. After summarising the current state-of-the-art in terms of taggers built from perturbative QCD, we moved on to concepts from machine learning; namely using so-called 'deep learning' neural networks to improve the performance of boosted top reconstruction, by building an image out of the calorimeter plane in a hadronic top quark event, and training image classifier algorithms over top signal against QCD background.

Quantifying its performance in terms of ROC curves (signal efficiency versus background rejection) we found comparable performance to well-established QCD based taggers, even in the presence of experimental degradation such as detector smearing and calorimeter energy cutoffs. The overarching question remains of whether these techniques will bear fruit when applied to real data and not just Monte Carlo simulation, but the robust performance of our network suggests that this is certainly a question that merits future investigation.

There are several other well-motivated directions for future work. For instance, in the global fit presented in chapter 3, a total of 12 parameters were constrained. However, in order to perform this fit, this parameter space was broken up into subsets of operators: 6 in

$t\bar{t}$ production, 3 in single top, and a further 3 in $t\bar{t}Z$ production. This factorisation, while easing the burden on the computational complexity of the fit, is not necessarily physically justified, since $t\bar{t}Z$ and other electroweak processes are sensitive to all 12 operators simultaneously, and because the $D = 6$ operators also affect top quark decays, the division of the fit into production and decay observables is an imperfect approximation. This can be improved by only fitting to fiducial top measurements that are presented in terms of final state quantities and not 'unfolded' to the level of tops. The experimental collaborations are beginning to favour presenting data in this way, and work on improving the capabilities of TopFitter to harness this data is underway.

Improvements can be made on the theory side as well. All the constraints presented here are at leading order in the SMEFT, and the bounds on the Wilson coefficients can be interpreted as valid at the scale that they are probed. However, just as in QCD, when we truncate the perturbative expansion in c_i/Λ^2 we introduce a scale dependence of the $D = 6$ Wilson coefficients. This can be modelled with an additional scale uncertainty propagated into each observable, or explicitly calculated through the RGEs for the $D = 6$ operators, which are known at 1-loop. RGE improvement cannot completely capture the full NLO corrections, however, because the new operators will in general induce additional loop corrections that can substantially affect the shapes of differential distributions. This is also related to the imperfectness of our modelling of (N)NLO corrections with QCD K-factors, which inevitably misses out on some kinematic effects. On the other hand, a complete 1-loop EFT calculation of $t\bar{t}$ production, for instance, is a formidable challenge, due to the sheer number of additional operators and diagrams involved, but certainly a worthwhile one, especially given the timescale of the future LHC programme.

Zooming out from the plethora of numerical results in this thesis, we can ask what general conclusions can be drawn? Firstly, though we often hear repeated that 'top quark physics has entered a precision era' and that 'the LHC is a top quark factory', it appears we still have some way to go before we saturate our understanding of the top quark's properties, and improvements in both experimental precision and theory understanding are both essential to this endeavour. Secondly, since at the time of writing there are no convincing hints of new resonant states from the current data and we are beginning to asymptote towards the maximum LHC reach for these states, it seems that precision understanding of the Standard Model, both as a full and an effective theory, will increasingly play a role in the hunt for new physics. It remains to be seen whether the $D = 6$ extension of the SM will be the avenue that leads us to the next Standard Model, but its usefulness as a tool for collider phenomenology has shown us beyond question that it is an avenue worth pursuing.

A $D = 6$ redefinitions of Standard Model input parameters

As well as generating additional contributions to S-matrix elements through new Feynman rules, the $D = 6$ operators also lead to modifications of the Standard Model Lagrangian parameters, which will propagate into observables. To see this, we note that an observable $\mathcal{O}$ such as a decay width or cross-section can be written as a function of Lagrangian parameters $\mathcal{O}(\{\rho\})$, where $\{\rho\} = \{g, g', v, y_t\}$ etc. These Lagrangian parameters bear a specific relation to the physical observables $\{obs\} = \{\alpha, G_F, m_Z, m_t\}$ so that the shift ϵ in $\mathcal{O}$ due to a Wilson coefficient c_i has, in addition to an explicit dependence on c_i, an implicit dependence due to a modification of the relation between Lagrangian parameters $\{\rho\}$ and observables $\{obs\}$:

$$\mathcal{O}(\{\rho(obs)\}) \to \mathcal{O} + \epsilon = \mathcal{O}(\{\rho(obs, c_i)\}, c_i) \tag{A.1}$$

In order to work consistently up to a given order in the EFT expansion, one must take care to ensure that the SM input parameters are appropriately renormalised as as function of c_i so that *obs* are unchanged. The modifications to the relations $\rho(obs)$ is the subject of this section[*].

Higgs potential

The Higgs potential receives a contribution from the operator $\mathcal{O}_\varphi$:

$$V(\varphi) = \lambda \left(\varphi^\dagger \varphi - \frac{1}{2}v^2 \right)^2 - c_\varphi(\varphi^\dagger \varphi)^3. \tag{A.2}$$

The shifted vacuum expectation value of the potential is at

$$\frac{1}{2}v_6^2 = \frac{1}{3c_\varphi}\left(\lambda - \lambda\sqrt{1 - \frac{3c_\varphi v^2}{\lambda}} \right), \tag{A.3}$$

This expression does not have a well-defined SM ($c_\varphi = 0$) limit, but this can be obtained by expanding to first order in c_φ:

$$\langle \varphi^\dagger \varphi \rangle = \frac{1}{2}v_6^2 = \frac{1}{2}v^2\left(1 + \frac{3c_\varphi v^2}{4\lambda} \right), \tag{A.4}$$

so that the shift in the vev is linear in c_φ and vanishes in the SM limit.

---

[*]To avoid unnecessary cluttering of notation, all $D = 6$ Wilson coefficients are dimensional, i.e. $c_i/\Lambda^2 \to c_i$ in this section.

Kinetic terms

The scalar kinetic part of the SM is modified by the operators $\mathcal{O}_{\varphi\square}$ and $\mathcal{O}_{\varphi D}$

$$\mathcal{L} = (D_\mu \varphi)^\dagger (D_\mu \varphi) + c_{\varphi\square}(\varphi^\dagger \varphi)\square(\varphi^\dagger \varphi) + c_{\varphi D}(\varphi^\dagger D_\mu \varphi)^*(\varphi^\dagger D^\mu \varphi), \tag{A.5}$$

In the unitary gauge, we can write the field φ as

$$\varphi = \begin{pmatrix} -i(1+\kappa_\pm)G_\pm \\ \frac{1}{\sqrt{2}}(v_6 + (1+\kappa_h)h + i(1+\kappa_0)G_0) \end{pmatrix}, \tag{A.6}$$

where the coefficients κ_h, κ_0 and $\kappa_\pm$ can be chosen to ensure the kinetic terms will be canonically normalised. Upon expanding out the Lagrangian and keeping just the scalar-only terms (the gauge parts will be treated shortly), we see that the kinetic terms are given by:

$$\begin{aligned} \mathcal{L}_\text{H} &= \frac{1}{2}(1+\kappa_h)^2 \left(1 + 2v^2\left[\frac{c_{\varphi D}}{4} - c_{\varphi\square}\right]\right)\partial_\mu h\partial^\mu h, \\ \mathcal{L}_\text{G0} &= \frac{1}{2}(1+\kappa_0)^2 \left(1 + 2v^2\frac{c_{\varphi D}}{4}\right)\partial_\mu G_0\partial^\mu G_0, \\ \mathcal{L}_\text{G\pm} &= \frac{1}{2}(1+\kappa_\pm)^2\partial_\mu G_+\partial^\mu G_- \end{aligned} \tag{A.7}$$

Canonical renormalisation of the kinetic terms then requires

$$\begin{aligned} \kappa_h &= v^2\left(c_{\varphi\square} - \frac{c_{\varphi D}}{4}\right), \\ \kappa_0 &= -v^2\frac{c_{\varphi D}}{4}, \\ \kappa_\pm &= 0. \end{aligned} \tag{A.8}$$

Combining the Higgs kinetic terms with the modified Higgs potential, the scalar Lagrangian reads

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}(\partial_\mu h)^2 - \frac{\kappa_h}{v_6^2}\left[h^2(\partial_\mu h)^2 + 2vh(\partial_\mu h)^2\right] - \lambda v_6^2\left(1 - \frac{3c_\varphi v^2}{2\lambda} + 2\kappa_h\right)h^2 \\ &\quad - \lambda v_6^2\left(1 - \frac{5c_\varphi v^2}{2\lambda} + 3\kappa_h\right)h^3 - \frac{1}{4}\lambda\left(1 - \frac{15c_\varphi v^2}{2\lambda} + 4\kappa_h\right)h^4 + \frac{3}{4}c_\varphi v h^5 + \frac{1}{8}c_\varphi h^6, \end{aligned} \tag{A.9}$$

so we see that there is a shift in the Higgs mass definition

$$m_h^2 = 2\lambda v_6^2\left(1 - \frac{3c_\varphi v^2}{2\lambda} + 2\kappa_h\right), \tag{A.10}$$

which does not correspond to a physical mass shift, as it can be absorbed into a renormal-isation of the quartic coupling $\lambda(m_h)$.

Gauge sector

The $D = 6$ operators induce redefinitions of the gauge fields and gauge couplings. The part of the $D = 6$ Lagrangian relevant for this discussion is:

$$
\begin{aligned}
\mathcal{L} \supset{}& c_{\varphi G}(\varphi^\dagger \varphi)G^A_{\mu\nu}G^{A,\mu\nu} + c_{\varphi W}(\varphi^\dagger \varphi)W^I_{\mu\nu}W^{I,\mu\nu} + c_{\varphi B}(\varphi^\dagger \varphi)B_{\mu\nu}B^{\mu\nu} \\
&+ c_{\varphi W B}(\varphi^\dagger \tau^I \varphi)W^I_{\mu\nu}B^{\mu\nu} + c_G f^{ABC}G^{A\nu}_\mu G^{B\nu}_\rho G^{C\rho}_\mu + c_W \epsilon^{IJK}W^{I\nu}_\mu W^{J\nu}_\rho W^{K\rho}_\mu .
\end{aligned}
\tag{A.11}
$$

After electroweak symmetry breaking, this becomes (combined with the $D = 4$ Lagrangian)

$$
\begin{aligned}
\mathcal{L} \supset{}& -\frac{1}{2}W^+_{\mu\nu}W^{\mu\nu}_- - \frac{1}{4}W^3_{\mu\nu}W^{\mu\nu}_3 - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}G^A_{\mu\nu}G^{A,\mu\nu} + \frac{1}{2}v_6^2 c_{\varphi G}G^A_{\mu\nu}G^{A,\mu\nu} \\
&+ \frac{1}{2}v_6^2 c_{\varphi W}W^I_{\mu\nu}W^{I,\mu\nu} + \frac{1}{2}v_6^2 c_{\varphi B}B_{\mu\nu}B^{\mu\nu} - \frac{1}{2}v_6^2 c_{\varphi W B}W^3_{\mu\nu}B^{\mu\nu}.
\end{aligned}
\tag{A.12}
$$

The gauge kinetic terms are no longer canonically normalised, and we have also induced kinetic mixing between W^3 and B. Beginning with the gluons, we can write the canonically normalised gluon field as $\mathcal{G}_\mu = (1 + \kappa_G)^{\frac{1}{2}}G_\mu$, then the Lagrangian reads

$$
\begin{aligned}
\mathcal{L} \supset{}& -\frac{1}{4}(1 - 2v_6^2 c_{\varphi G})G_{\mu\nu}G^{\mu\nu} \\
={}& -\frac{1}{4}\mathcal{G}_{\mu\nu}\mathcal{G}^{\mu\nu},
\end{aligned}
\tag{A.13}
$$

provided that

$$
\kappa_G = -2v_6^2 c_{\varphi G}.
\tag{A.14}
$$

The gluon field is rescaled

$$
G^A_\mu = \mathcal{G}^A_\mu(1 + c_{\varphi G}v^2),
\tag{A.15}
$$

so that the strong coupling constant must also be renormalised

$$
\bar{g}_s = g_s(1 + c_{\varphi G}v^2),
\tag{A.16}
$$

in order to keep the vector currents unchanged, i.e. $g_s G^A_\mu = \bar{g}_s \mathcal{G}^A_\mu$. The exercise can be repeated for the electroweak bosons

$$
\begin{aligned}
W^I_\mu &= \mathcal{W}^I_\mu(1 + c_{\varphi W}v^2) \\
B^\mu &= \mathcal{B}_\mu(1 + c_{\varphi B}v^2)
\end{aligned}
\tag{A.17}
$$

where the renormalisation coefficients are $\kappa_W = -2v^2 c_{\varphi W}$ and $\kappa_B = -2v^2 c_{\varphi B}$, and the gauge couplings are renormalised to

$$\begin{aligned}
\bar{g} &= g(1 + c_{\varphi W} v^2), \\
\bar{g}' &= g'(1 + c_{\varphi B} v^2).
\end{aligned} \tag{A.18}$$

We must also take into account the kinetic mixing that has been induced. The Lagrangian for the kinetic terms of the $\mathcal{W}^3$ and $\mathcal{B}$ fields can be written as a non-diagonal matrix:

$$\mathcal{L} = -\frac{1}{4} \begin{bmatrix} \mathcal{W}^3_{\mu\nu} & \mathcal{B}_{\mu\nu} \end{bmatrix} \begin{bmatrix} 1 & v_6^2 c_{\varphi W B} \\ v_6^2 c_{\varphi W B} & 1 \end{bmatrix} \begin{bmatrix} \mathcal{W}^3_{\mu\nu} \\ \mathcal{B}_{\mu\nu} \end{bmatrix}. \tag{A.19}$$

Canonically normalising the $\mathcal{W}^3$ and $\mathcal{B}$ fields then just amounts to diagonalising this matrix, which can be done by rotating the fields

$$\begin{bmatrix} \mathcal{W}^3_{\mu} \\ \mathcal{B}_{\mu} \end{bmatrix} = \begin{bmatrix} 1 & v_6^2 c_{\varphi W B} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathcal{W}^3_{\mu} \\ \mathcal{B}_{\mu} \end{bmatrix}. \tag{A.20}$$

The canonically normalised kinetic Lagrangian is then

$$\mathcal{L} = -\frac{1}{4} \mathcal{W}^I_{\mu\nu} \mathcal{W}^{\mu\nu}_I - \frac{1}{4} \mathcal{B}_{\mu\nu} \mathcal{B}^{\mu\nu} = -\frac{1}{2} \mathcal{W}^+_{\mu\nu} \mathcal{W}^{\mu\nu}_- - \frac{1}{4} \mathcal{W}^3_{\mu\nu} \mathcal{W}^{\mu\nu}_3 - \frac{1}{4} \mathcal{B}_{\mu\nu} \mathcal{B}^{\mu\nu}, \tag{A.21}$$

where $\mathcal{W}^{1,2}_{\mu} = \mathcal{W}^{1,2}_{\mu}$.

Turning now to the electroweak mass terms, we start with the Lagrangian

$$\mathcal{L} = -\frac{1}{4} \bar{g}^2 v_6^2 \mathcal{W}^+_{\mu} \mathcal{W}^{\mu}_- + \frac{1}{8} v_6^2 (\bar{g} \mathcal{W}^3_{\mu} - \bar{g}' \mathcal{B}_{\mu})^2 + \frac{1}{16} v_6^2 c_{\varphi D} (\bar{g} \mathcal{W}^3_{\mu} - \bar{g}' \mathcal{B}_{\mu})^2, \tag{A.22}$$

so that the W mass term can be straightforwardly read off as

$$M_W^2 = \frac{1}{4} \bar{g}^2 v_6^2. \tag{A.23}$$

As for the Z and photon, we can write the Lagrangian as

$$\mathcal{L} = \frac{1}{2} \left(\frac{1}{4} v_6^2 \left\{ 1 + \frac{1}{2} v_6^2 c_{\varphi D} \right\} \right) \begin{bmatrix} \mathcal{W}^3_{\mu\nu} & \mathcal{B}_{\mu\nu} \end{bmatrix} \begin{bmatrix} \bar{g}^2 & \bar{g}\bar{g}' \\ \bar{g}\bar{g}' & \bar{g}'^2 \end{bmatrix} \begin{bmatrix} \mathcal{W}^3_{\mu} \\ \mathcal{B}_{\mu} \end{bmatrix} \tag{A.24}$$

Diagonalising the mass matrix gives us the updated expression

$$\begin{bmatrix} \mathcal{Z}_{\mu} \\ \mathcal{A}_{\mu} \end{bmatrix} = \begin{bmatrix} \cos\bar{\theta}_W & -\sin\bar{\theta}_W \\ \sin\bar{\theta}_W & \cos\bar{\theta}_W \end{bmatrix} \begin{bmatrix} \mathcal{W}^3_{\mu} \\ \mathcal{B}_{\mu} \end{bmatrix}, \tag{A.25}$$

where

$$\cos \bar{\theta}_W = \frac{\bar{g}}{\sqrt{\bar{g}^2 + \bar{g}'^2}} \left[1 + v^2 \frac{\bar{g}}{\bar{g}'} \frac{\bar{g}'^2}{\bar{g}^2 + \bar{g}'^2} c_{\varphi WB} \right]$$

$$\sin \bar{\theta}_W = \frac{\bar{g}'}{\sqrt{\bar{g}^2 + \bar{g}'^2}} \left[1 - v^2 \frac{\bar{g}'}{\bar{g}} \frac{\bar{g}^2}{\bar{g}^2 + \bar{g}'^2} c_{\varphi WB} \right].$$

(A.26)

Then the Z mass is given by

$$M_Z^2 = \frac{v_6^2}{4}(\bar{g}^2 + \bar{g}'^2) + \frac{1}{8} v_6^4 c_{\varphi D}(\bar{g}^2 + \bar{g}'^2) + \frac{1}{2} v_6^4 \bar{g}\bar{g}' c_{\varphi WB},$$

(A.27)

while the photon remains massless as required. Substituting the renormalised gauge fields into the covariant derivative operator gives the expression

$$D_\mu = \partial_\mu + i \frac{g'}{\sqrt{2}} [\boldsymbol{W}_\mu^+ T^+ + \boldsymbol{W}_\mu^- T^-] + i \bar{g}_Z [T^3 - (\sin^2 \bar{\theta} + \kappa)Q] \boldsymbol{Z}_\mu + i \bar{e} Q \boldsymbol{A}_\mu,$$

(A.28)

where, as usual, the generators $T^\pm = T_1 \mp i T_2$ and the electric charge is $Q = T_3 + Y$. The effective neutral couplings are

$$\bar{e} = \frac{\bar{g}\bar{g}'}{\sqrt{\bar{g}^2 + \bar{g}'^2}} \left[1 - \frac{\bar{g}\bar{g}'}{g^2 + g'^2} v_6^2 c_{\varphi WB} \right] = \bar{g} \sin \bar{\theta} - \frac{1}{2} \cos \bar{\theta} \bar{g} v_6^2 c_{\varphi WB}$$

$$\bar{g}_Z = \sqrt{\bar{g}^2 + \bar{g}'^2} + \frac{\bar{g}\bar{g}'}{g^2 + g'^2} v_6^2 c_{\varphi WB} = \frac{\bar{e}}{\sin \bar{\theta} \cos \bar{\theta}} \left[1 + \frac{\bar{g}^2 + \bar{g}'^2}{2\bar{g}\bar{g}'} v_6^2 c_{\varphi WB} \right]$$

$$\sin^2 \bar{\theta} = \frac{\bar{g}'^2}{\bar{g}^2 + \bar{g}'^2} + \frac{\bar{g}\bar{g}'(\bar{g}^2 - \bar{g}'^2)}{(\bar{g}^2 + \bar{g}'^2)^2} v_6^2 c_{\varphi WB}.$$

(A.29)

Yukawa sector

The Yukawa sector will be modified by the operators of type $\psi^2 \varphi^3$. The Yukawa Lagrangian for the unbroken theory now reads

$$\mathcal{L}_{\text{Yukawa}} = -(\varphi^\dagger \bar{d}_s [y_d]_{st} Q_t + \tilde{\varphi}^\dagger \bar{u}_s [y_u]_{st} Q_t + \varphi^\dagger \bar{e}_s [y_e]_{st} L_t + h.c.)$$
$$+ (c_{d\varphi}(\varphi^\dagger \varphi)\varphi^\dagger \bar{d}_s Q_s + c_{u\varphi}(\varphi^\dagger \varphi)\varphi^\dagger \bar{u}_s Q_s + c_{u\varphi}(\varphi^\dagger \varphi)\varphi^\dagger \bar{e}_s L_s + h.c.),$$

(A.30)

which in the unbroken theory leads to the fermion mass matrices

$$[M_\psi]_{rs} = \frac{v_6}{\sqrt{2}} \left([y_\psi]_{rs} - \frac{1}{2} v^2 [c_{\psi\varphi}]_{rs} \right), \quad \text{where} \quad \psi = u, d, e$$

(A.31)

and to the Higgs fermion couplings

$$[\kappa_\psi]_{rs} = \frac{1}{\sqrt{2}}[y_\psi]_{rs}(1 + \kappa_h) - \frac{3}{2\sqrt{2}}v^2 c_{\psi\varphi}$$
$$= \frac{1}{\sqrt{2}}[M_\psi]_{rs}(1 + \kappa_h) - \frac{v^2}{\sqrt{2}}v^2 c_{\psi\varphi}, \quad \text{where} \quad \psi = u, d, e \tag{A.32}$$

which, unlike in the SM, are not simply proportional to the fermion mass matrices. Furthermore, because the fermion mass matrices and Yukawa matrices have different RGEs, they are not simultaneously diagonalisable, so Higgs-fermion couplings will be no longer flavour diagonal.

Fermi sector

The Fermi coupling constant is measured from the transition rate for $\mu^- \to e^- \bar{\nu}_e \nu_\mu$, and this in turn defines the value of the electroweak scale v. In the SM alone, this is described by the effective operator

$$\mathcal{L}_{G_F} = \frac{4G_F}{\sqrt{2}}(\bar{\nu}_\mu \gamma^\mu P_L \mu)(\bar{e}\gamma_\mu P_L \nu_e). \tag{A.33}$$

The Fermi constant G_F will also receive corrections from $D = 6$ operators, leading to the new value:

$$\sqrt{2}G_F = \frac{1}{v_6^2} - \frac{1}{2}\left(c_{ll}^{2112} + c_{ll}^{1221}\right) + \left(c_{\varphi l}^{(3)11} + c_{\varphi l}^{(3)22}\right). \tag{A.34}$$

Although this looks like a physical shift, in fact it can be combined with the modified expression for the Higgs mass to define renormalised values for the Higgs self-coupling and vev:

$$\lambda = \frac{3\sqrt{2}c_\varphi}{4G_F} + \frac{m_h^2}{4}\left(c_{ll}^{2112} + c_{ll}^{1221} - 2(c_{\varphi l}^{(3)11} + c_{\varphi l}^{(3)22})\right) + \frac{1}{\sqrt{2}}m_h^2 G_F(1 - 2\kappa_h)$$
$$v_6 = \frac{1}{(\sqrt{2}G_F)^{\frac{1}{2}}} + \frac{1}{2(\sqrt{2}G_F)^{\frac{3}{2}}}\left(c_{\varphi l}^{(3)11} + c_{\varphi l}^{(3)22} - \frac{1}{2}\left(c_{ll}^{2112} + c_{ll}^{1221}\right)\right). \tag{A.35}$$

Likewise, we can rearrange the expressions for the observables $\alpha_{em}(\equiv e^2/4\pi)$ and M_Z to obtain the renormalised $U(1)_Y$ and $SU(2)_L$ couplings

$$\bar{g}' = g' + \frac{v_6^2(4c_{\varphi WB}g + c_{\varphi D}g')(-4M_Z^2 + v_6^2(g^2 - g'^2 + 16\pi\alpha_{em}))}{32(M_Z^2 - 4\pi v_6^2 \alpha_{em})}$$
$$\bar{g} = g - \frac{2v_6^2(4c_{\varphi WB}g + c_{\varphi D}g')(4M_Z^4 + M_Z^2 v_6^2)}{M_Z^2(M_Z^2 - 4\pi v_6^2 \alpha_{em})}$$
$$\times \left[\frac{v_6^4((g^2 - g'^2 - 20\pi\alpha_{em}) + \pi\alpha_{em}(-3g^2 + 3g'^2 + 16\pi\alpha_{em}))}{(4M_Z^2 + v_6^2(g^2 - g'^2))^2}\right] \tag{A.36}$$

where the SM coupling constants are as usual given by

$$
\begin{aligned}
g' &= \frac{\sqrt{2}}{v_6} \left(M_Z^2 - \sqrt{M_Z^4 - 4\pi\alpha_{em} M_Z^2 v_6^2} \right)^{\frac{1}{2}} \\
g &= \frac{\sqrt{2}}{v_6} \left(M_Z^2 + \sqrt{M_Z^4 - 4\pi\alpha_{em} M_Z^2 v_6^2} \right)^{\frac{1}{2}}.
\end{aligned}
\tag{A.37}
$$

This completes the finite renormalisation of the Standard Model Lagrangian due to the effects of operators of dimension $D = 6$. The modifications to all other derived parameters can be obtained from the relations presented here.

List of Figures

List of Tables

References

[1] A. Buckley, C. Englert, J. Ferrando, D. J. Miller, L. Moore, M. Russell, and C. D. White, "Global fit of top quark effective theory to data," *Phys. Rev.* **D92** no. 9, (2015) 091501, `arXiv:1506.08845 [hep-ph]`.

[2] A. Buckley, C. Englert, J. Ferrando, D. J. Miller, L. Moore, M. Russell, and C. D. White, "Constraining top quark effective theory in the LHC Run II era," *JHEP* **04** (2016) 015, `arXiv:1512.03360 [hep-ph]`.

[3] C. Englert, L. Moore, K. Nordström, and M. Russell, "Giving top quark effective operators a boost," *Phys. Lett.* **B763** (2016) 9–15, `arXiv:1607.04304 [hep-ph]`.

[4] C. Englert and M. Russell, "Top quark electroweak couplings at future lepton colliders," `arXiv:1704.01782 [hep-ph]`.

[5] G. Kasieczka, T. Plehn, M. Russell, and T. Schell, "Deep-learning Top Taggers or The End of QCD?," *JHEP* **05** (2017) 006, `arXiv:1701.08784 [hep-ph]`.

[6] H. Yukawa, "On the Interaction of Elementary Particles I," *Proc. Phys. Math. Soc. Jap.* **17** (1935) 48–57. [Prog. Theor. Phys. Suppl.1,1(1935)].

[7] C. M. G. Lattes, H. Muirhead, G. P. S. Occhialini, and C. F. Powell, "PROCESSES INVOLVING CHARGED MESONS," *Nature* **159** (1947) 694–697. [,42(1947)].

[8] J. Chadwick, "Possible Existence of a Neutron," *Nature* **129** (1932) 312.

[9] E. Fermi, "An attempt of a theory of beta radiation. 1.," *Z. Phys.* **88** (1934) 161–177.

[10] R. P. Feynman and M. Gell-Mann, "Theory of Fermi interaction," *Phys. Rev.* **109** (1958) 193–198.

[11] E. C. G. Sudarshan and R. e. Marshak, "Chirality invariance and the universal Fermi interaction," *Phys. Rev.* **109** (1958) 1860–1860.

[12] J. S. Schwinger, "A Theory of the Fundamental Interactions," *Annals Phys.* **2** (1957) 407–434.

[13] T. D. Lee and C.-N. Yang, "Possible Nonlocal Effects in μ Decay," *Phys. Rev.* **108** (1957) 1611–1614.

[14] S. A. Bludman, "On the universal Fermi interaction," *Nuovo Cim.* **9** (1958) 433–445.

[15] J. Leite Lopes, "A model of the universal fermi interaction," *Nucl. Phys.* **8** (1958) 234–236.

[16] T. D. Lee and C.-N. Yang, "Question of Parity Conservation in Weak Interactions," *Phys. Rev.* **104** (1956) 254–258.

[17] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, "Experimental Test of Parity Conservation in Beta Decay," *Phys. Rev.* **105** (1957) 1413–1414.

[18] G. Zweig, "An SU(3) model for strong interaction symmetry and its breaking. Version 2," in *DEVELOPMENTS IN THE QUARK THEORY OF HADRONS. VOL. 1. 1964 - 1978*, D. Lichtenberg and S. P. Rosen, eds., pp. 22–101. 1964. http://inspirehep.net/record/4674/files/cern-th-412.pdf.

[19] M. Gell-Mann, "A Schematic Model of Baryons and Mesons," *Phys. Lett.* **8** (1964) 214–215.

[20] C.-N. Yang and R. L. Mills, "Conservation of Isotopic Spin and Isotopic Gauge Invariance," *Phys. Rev.* **96** (1954) 191–195.

[21] M. J. G. Veltman, "Perturbation theory of massive Yang-Mills fields," *Nucl. Phys.* **B7** (1968) 637–650.

[22] M. J. G. Veltman, "Generalized ward identities and yang-mills fields," *Nucl. Phys.* **B21** (1970) 288–302.

[23] P. W. Anderson, "Plasmons, Gauge Invariance, and Mass," *Phys. Rev.* **130** (1963) 439–442.

[24] J. Goldstone, "Field Theories with Superconductor Solutions," *Nuovo Cim.* **19** (1961) 154–164.

[25] J. Goldstone, A. Salam, and S. Weinberg, "Broken Symmetries," *Phys. Rev.* **127** (1962) 965–970.

[26] P. W. Higgs, "Broken symmetries, massless particles and gauge fields," *Phys. Lett.* **12** (1964) 132–133.

[27] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.* **13** (1964) 508–509.

[28] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys. Rev. Lett.* **13** (1964) 321–323.

[29] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles," *Phys. Rev. Lett.* **13** (1964) 585–587.

[30] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett.* **19** (1967) 1264–1266.

[31] A. Salam, "Weak and Electromagnetic Interactions," *Conf. Proc.* **C680519** (1968) 367–377.

[32] S. L. Glashow, "Partial Symmetries of Weak Interactions," *Nucl. Phys.* **22** (1961) 579–588.

[33] A. Salam and J. C. Ward, "Electromagnetic and weak interactions," *Phys. Lett.* **13** (1964) 168–171.

[34] R. P. Feynman, "Space-time approach to nonrelativistic quantum mechanics," *Rev. Mod. Phys.* **20** (1948) 367–387.

[35] R. P. Feynman, "Space - time approach to quantum electrodynamics," *Phys. Rev.* **76** (1949) 769–789.

[36] R. P. Feynman, "Mathematical formulation of the quantum theory of electromagnetic interaction," *Phys. Rev.* **80** (1950) 440–457.

[37] J. S. Schwinger, "Quantum electrodynamics. I A covariant formulation," *Phys. Rev.* **74** (1948) 1439.

[38] J. S. Schwinger, "Quantum electrodynamics. 2. Vacuum polarization and selfenergy," *Phys. Rev.* **75** (1948) 651.

[39] J. S. Schwinger, "Quantum electrodynamics. III: The electromagnetic properties of the electron: Radiative corrections to scattering," *Phys. Rev.* **76** (1949) 790–817.

[40] S. Tomonaga, "On a relativistically invariant formulation of the quantum theory of wave fields," *Prog. Theor. Phys.* **1** (1946) 27–42.

[41] Z. Koba, T. Tati, and S. i. Tomonaga, "On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields. II: Case of Interacting Electromagnetic and Electron Fields," *Prog. Theor. Phys.* **2** no. 3, (1947) 101–116.

[42] F. J. Dyson, "The S matrix in quantum electrodynamics," *Phys. Rev.* **75** (1949) 1736–1755.

[43] F. J. Dyson, "Divergence of perturbation theory in quantum electrodynamics," *Phys. Rev.* **85** (1952) 631–632.

[44] F. J. Dyson, "The Radiation theories of Tomonaga, Schwinger, and Feynman," *Phys. Rev.* **75** (1949) 486–502.

[45] G. 't Hooft and M. J. G. Veltman, "Regularization and Renormalization of Gauge Fields," *Nucl. Phys.* **B44** (1972) 189–213.

[46] G. 't Hooft, "Renormalization of Massless Yang-Mills Fields," *Nucl. Phys.* **B33** (1971) 173–199.

[47] G. 't Hooft, "Renormalizable Lagrangians for Massive Yang-Mills Fields," *Nucl. Phys.* **B35** (1971) 167–188.

[48] E. D. Bloom *et al.*, "High-Energy Inelastic e p Scattering at 6-Degrees and 10-Degrees," *Phys. Rev. Lett.* **23** (1969) 930–934.

[49] M. Breidenbach, J. I. Friedman, H. W. Kendall, E. D. Bloom, D. H. Coward, H. C. DeStaebler, J. Drees, L. W. Mo, and R. E. Taylor, "Observed Behavior of Highly Inelastic electron-Proton Scattering," *Phys. Rev. Lett.* **23** (1969) 935–939.

[50] J. D. Bjorken, "Asymptotic Sum Rules at Infinite Momentum," *Phys. Rev.* **179** (1969) 1547–1553.

[51] R. P. Feynman, "Very high-energy collisions of hadrons," *Phys. Rev. Lett.* **23** (1969) 1415–1417.

[52] D. J. Gross and F. Wilczek, "Ultraviolet Behavior of Nonabelian Gauge Theories," *Phys. Rev. Lett.* **30** (1973) 1343–1346.

[53] H. D. Politzer, "Reliable Perturbative Results for Strong Interactions?," *Phys. Rev. Lett.* **30** (1973) 1346–1349.

[54] S. L. Glashow, J. Iliopoulos, and L. Maiani, "Weak Interactions with Lepton-Hadron Symmetry," *Phys. Rev.* **D2** (1970) 1285–1292.

[55] **SLAC-SP-017** Collaboration, J. E. Augustin *et al.*, "Discovery of a Narrow Resonance in e^+e^- Annihilation," *Phys. Rev. Lett.* **33** (1974) 1406–1408. [Adv. Exp. Phys.5,141(1976)].

[56] **E598** Collaboration, J. J. Aubert *et al.*, "Experimental Observation of a Heavy Particle J," *Phys. Rev. Lett.* **33** (1974) 1404–1406.

[57] M. Kobayashi and T. Maskawa, "CP Violation in the Renormalizable Theory of Weak Interaction," *Prog. Theor. Phys.* **49** (1973) 652–657.

[58] N. Cabibbo, "Unitary Symmetry and Leptonic Decays," *Phys. Rev. Lett.* **10** (1963) 531–533. [,648(1963)].

[59] S. W. Herb *et al.*, "Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions," *Phys. Rev. Lett.* **39** (1977) 252–255.

[60] W. R. Innes *et al.*, "Observation of structure in the Υ region," *Phys. Rev. Lett.* **39** (1977) 1240–1242. [Erratum: Phys. Rev. Lett.39,1640(1977)].

[61] **CDF** Collaboration, F. Abe *et al.*, "Observation of top quark production in $\bar{p}p$ collisions," *Phys. Rev. Lett.* **74** (1995) 2626–2631, `arXiv:hep-ex/9503002 [hep-ex]`.

[62] **D0** Collaboration, S. Abachi *et al.*, "Observation of the top quark," *Phys. Rev. Lett.* **74** (1995) 2632–2637, `arXiv:hep-ex/9503003 [hep-ex]`.

[63] M. L. Perl *et al.*, "Evidence for Anomalous Lepton Production in e+ - e- Annihilation," *Phys. Rev. Lett.* **35** (1975) 1489–1492.

[64] **DONUT** Collaboration, K. Kodama *et al.*, "Observation of tau neutrino interactions," *Phys. Lett.* **B504** (2001) 218–224, `arXiv:hep-ex/0012035 [hep-ex]`.

[65] **Gargamelle Neutrino** Collaboration, F. J. Hasert *et al.*, "Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment," *Phys. Lett.* **46B** (1973) 138–140.

[66] F. J. Hasert *et al.*, "Search for Elastic ν_μ Electron Scattering," *Phys. Lett.* **46B** (1973) 121–124.

[67] **UA1** Collaboration, G. Arnison *et al.*, "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV," *Phys. Lett.* **122B** (1983) 103–116. [,611(1983)].

[68] **UA1** Collaboration, G. Arnison *et al.*, "Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider," *Phys. Lett.* **126B** (1983) 398–410.

[69] **UA2** Collaboration, M. Banner *et al.*, "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider," *Phys. Lett.* **122B** (1983) 476–485.

[70] **UA2** Collaboration, P. Bagnaia *et al.*, "Evidence for Z0 —¿ e+ e- at the CERN anti-p p Collider," *Phys. Lett.* **B129** (1983) 130–140.

[71] **TASSO** Collaboration, R. Brandelik *et al.*, "Evidence for Planar Events in e+ e- Annihilation at High-Energies," *Phys. Lett.* **B86** (1979) 243–249.

[72] **SLD Electroweak Group, DELPHI, ALEPH, SLD, SLD Heavy Flavour Group, OPAL, LEP Electroweak Working Group, L3** Collaboration, S. Schael *et al.*, "Precision electroweak measurements on the Z resonance," *Phys. Rept.* **427** (2006) 257–454, `arXiv:hep-ex/0509008 [hep-ex]`.

[73] **ATLAS** Collaboration, G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys. Lett.* **B716** (2012) 1–29, `arXiv:1207.7214 [hep-ex]`.

[74] **CMS** Collaboration, S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys. Lett.* **B716** (2012) 30–61, `arXiv:1207.7235 [hep-ex]`.

[75] **LHC Higgs Cross Section Working Group** Collaboration, D. de Florian *et al.*, "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector," `arXiv:1610.07922 [hep-ph]`.

[76] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, "Advantages of the Color Octet Gluon Picture," *Phys. Lett.* **B47** (1973) 365–368.

[77] *see e.g.* , M. D. Schwartz, *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014. `http://www.cambridge.org/us/academic/subjects/physics/theoretical-physics-and-mathematical-physics/quantum-field-theory-and-standard-model`.

[78] B. W. Lee, C. Quigg, and H. B. Thacker, "Weak Interactions at Very High-Energies: The Role of the Higgs Boson Mass," *Phys. Rev.* **D16** (1977) 1519.

[79] B. W. Lee, C. Quigg, and H. B. Thacker, "The Strength of Weak Interactions at Very High-Energies and the Higgs Boson Mass," *Phys. Rev. Lett.* **38** (1977) 883–885.

[80] A. Djouadi, "The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model," *Phys. Rept.* **457** (2008) 1–216, `arXiv:hep-ph/0503172 [hep-ph]`.

[81] P. Sikivie, L. Susskind, M. B. Voloshin, and V. I. Zakharov, "Isospin Breaking in Technicolor Models," *Nucl. Phys.* **B173** (1980) 189–207.

[82] C. H. Llewellyn Smith, "High-Energy Behavior and Gauge Symmetry," *Phys. Lett.* **B46** (1973) 233–236.

[83] J. S. Bell, "HIGH-ENERGY BEHAVIOR OF TREE DIAGRAMS IN GAUGE THEORIES," *Nucl. Phys.* **B60** (1973) 427–436.

[84] **Particle Data Group** Collaboration, K. A. Olive *et al.*, "Review of Particle Physics," *Chin. Phys.* **C38** (2014) 090001.

[85] L. Wolfenstein, "Parametrization of the Kobayashi-Maskawa Matrix," *Phys. Rev. Lett.* **51** (1983) 1945.

[86] B. Pontecorvo, "Inverse beta processes and nonconservation of lepton charge," *Sov. Phys. JETP* **7** (1958) 172–173. [Zh. Eksp. Teor. Fiz.34,247(1957)].

[87] Z. Maki, M. Nakagawa, and S. Sakata, "Remarks on the unified model of elementary particles," *Prog. Theor. Phys.* **28** (1962) 870–880.

[88] B. Pontecorvo, "Neutrino Experiments and the Problem of Conservation of Leptonic Charge," *Sov. Phys. JETP* **26** (1968) 984–988. [Zh. Eksp. Teor. Fiz.53,1717(1967)].

[89] **ZEUS, H1** Collaboration, F. D. Aaron *et al.*, "Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA," *JHEP* **01** (2010) 109, `arXiv:0911.0884 [hep-ex]`.

[90] **DELPHI, OPAL, LEP Electroweak, ALEPH, L3** Collaboration, S. Schael *et al.*, "Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP," *Phys. Rept.* **532** (2013) 119–244, `arXiv:1302.3415 [hep-ex]`.

[91] O. Brein, A. Djouadi, and R. Harlander, "NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders," *Phys. Lett.* **B579** (2004) 149–156, `arXiv:hep-ph/0307206 [hep-ph]`.

[92] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, "New parton distribution functions from a global analysis of quantum chromodynamics," *Phys. Rev.* **D93** no. 3, (2016) 033006, `arXiv:1506.07443 [hep-ph]`.

[93] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, "Parton distributions in the LHC era: MMHT 2014 PDFs," *Eur. Phys. J.* **C75** no. 5, (2015) 204, `arXiv:1412.3989 [hep-ph]`.

[94] **NNPDF** Collaboration, R. D. Ball *et al.*, "Parton distributions for the LHC Run II," *JHEP* **04** (2015) 040, `arXiv:1410.8849 [hep-ph]`.

[95] S. Alekhin, S. Alioli, R. D. Ball, V. Bertone, J. Blumlein, and others, "The PDF4LHC Working Group Interim Report," `arXiv:1101.0536 [hep-ph]`.

[96] A. Martin, W. Stirling, R. Thorne, and G. Watt, "Parton distributions for the LHC," *Eur.Phys.J.* **C63** (2009) 189–285, `arXiv:0901.0002 [hep-ph]`.

[97] P. A. Baikov, K. G. Chetyrkin, and J. H. Kühn, "Five-Loop Running of the QCD coupling constant," *Phys. Rev. Lett.* **118** no. 8, (2017) 082002, `arXiv:1606.08659 [hep-ph]`.

[98] T. Luthe, A. Maier, P. Marquard, and Y. Schroder, "Complete renormalization of QCD at five loops," *JHEP* **03** (2017) 020, `arXiv:1701.07068 [hep-ph]`.

[99] F. Herzog, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, "The five-loop beta function of Yang-Mills theory with fermions," *JHEP* **02** (2017) 090, `arXiv:1701.01404 [hep-ph]`.

[100] A. H. Hoang, "The Top Mass: Interpretation and Theoretical Uncertainties," in *Proceedings, 7th International Workshop on Top Quark Physics (TOP2014): Cannes, France, September 28-October 3, 2014.* 2014. `arXiv:1412.3649 [hep-ph]`. `https://inspirehep.net/record/1333866/files/arXiv:1412.3649.pdf`.

[101] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "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, `arXiv:1405.0301 [hep-ph]`.

[102] M. Czakon and A. Mitov, "NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels," *JHEP* **12** (2012) 054, `arXiv:1207.0236 [hep-ph]`.

[103] M. Czakon and A. Mitov, "NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction," *JHEP* **01** (2013) 080, `arXiv:1210.6832 [hep-ph]`.

[104] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$," *Phys. Rev. Lett.* **110** (2013) 252004, `arXiv:1303.6254 [hep-ph]`.

[105] J. M. Campbell, R. K. Ellis, P. Nason, and E. Re, "Top-Pair Production and Decay at NLO Matched with Parton Showers," *JHEP* **04** (2015) 114, `arXiv:1412.1828 [hep-ph]`.

[106] S. Frixione, F. Stoeckli, P. Torrielli, B. R. Webber, and C. D. White, "The MCaNLO 4.0 Event Generator," `arXiv:1010.0819 [hep-ph]`.

[107] J. M. Campbell and R. K. Ellis, "MCFM for the Tevatron and the LHC," *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 10–15, `arXiv:1007.3492 [hep-ph]`.

[108] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, "Event generation with SHERPA 1.1," *JHEP* **02** (2009) 007, `arXiv:0811.4622 [hep-ph]`.

[109] M. Czakon, D. Heymes, and A. Mitov, "High-precision differential predictions for top-quark pairs at the LHC," *Phys. Rev. Lett.* **116** no. 8, (2016) 082003, `arXiv:1511.00549 [hep-ph]`.

[110] M. Czakon, P. Fiedler, D. Heymes, and A. Mitov, "NNLO QCD predictions for fully-differential top-quark pair production at the Tevatron," *JHEP* **05** (2016) 034, `arXiv:1601.05375 [hep-ph]`.

[111] W. Bernreuther and Z.-G. Si, "Top quark and leptonic charge asymmetries for the Tevatron and LHC," *Phys. Rev.* **D86** (2012) 034026, `arXiv:1205.6580 [hep-ph]`.

[112] J. A. Aguilar-Saavedra, D. Amidei, A. Juste, and M. Perez-Victoria, "Asymmetries in top quark pair production at hadron colliders," *Rev. Mod. Phys.* **87** (2015) 421–455, `arXiv:1406.1798 [hep-ph]`.

[113] J. A. Aguilar-Saavedra and M. Perez-Victoria, "Probing the Tevatron t tbar asymmetry at LHC," *JHEP* **05** (2011) 034, `arXiv:1103.2765 [hep-ph]`.

[114] J. F. Kamenik, J. Shu, and J. Zupan, "Review of New Physics Effects in t-tbar Production," *Eur. Phys. J.* **C72** (2012) 2102, `arXiv:1107.5257 [hep-ph]`.

[115] J. H. Kuhn and G. Rodrigo, "Charge asymmetries of top quarks at hadron colliders revisited," *JHEP* **01** (2012) 063, `arXiv:1109.6830 [hep-ph]`.

[116] M. Czakon, P. Fiedler, and A. Mitov, "Resolving the Tevatron Top Quark Forward-Backward Asymmetry Puzzle: Fully Differential Next-to-Next-to-Leading-Order Calculation," *Phys. Rev. Lett.* **115** no. 5, (2015) 052001, `arXiv:1411.3007 [hep-ph]`.

[117] Q.-H. Cao, R. Schwienhorst, and C. P. Yuan, "Next-to-leading order corrections to single top quark production and decay at Tevatron. 1. s^- channel process," *Phys. Rev.* **D71** (2005) 054023, `arXiv:hep-ph/0409040 [hep-ph]`.

[118] J. M. Campbell, R. K. Ellis, and F. Tramontano, "Single top production and decay at next-to-leading order," *Phys. Rev.* **D70** (2004) 094012, `arXiv:hep-ph/0408158 [hep-ph]`.

[119] N. Kidonakis, "NNLL resummation for s-channel single top quark production," *Phys. Rev.* **D81** (2010) 054028, `arXiv:1001.5034 [hep-ph]`.

[120] N. Kidonakis, "Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production," *Phys. Rev.* **D83** (2011) 091503, `arXiv:1103.2792 [hep-ph]`.

[121] J. M. Campbell, R. Frederix, F. Maltoni, and F. Tramontano, "Next-to-Leading-Order Predictions for t-Channel Single-Top Production at Hadron Colliders," *Phys. Rev. Lett.* **102** (2009) 182003, `arXiv:0903.0005 [hep-ph]`.

[122] P. Falgari, P. Mellor, and A. Signer, "Production-decay interferences at NLO in QCD for t-channel single-top production," *Phys. Rev.* **D82** (2010) 054028, `arXiv:1007.0893 [hep-ph]`.

[123] Q.-H. Cao, R. Schwienhorst, J. A. Benitez, R. Brock, and C. P. Yuan, "Next-to-leading order corrections to single top quark production and decay at the Tevatron: 2. t^- channel process," *Phys. Rev.* **D72** (2005) 094027, `arXiv:hep-ph/0504230 [hep-ph]`.

[124] M. Brucherseifer, F. Caola, and K. Melnikov, "On the NNLO QCD corrections to single-top production at the LHC," *Phys. Lett.* **B736** (2014) 58–63, `arXiv:1404.7116 [hep-ph]`.

[125] S. S. D. Willenbrock and D. A. Dicus, "Production of Heavy Quarks from W Gluon Fusion," *Phys. Rev.* **D34** (1986) 155.

[126] T. M. P. Tait, "The tW^- mode of single top production," *Phys. Rev.* **D61** (1999) 034001, `arXiv:hep-ph/9909352 [hep-ph]`.

[127] S. Zhu, "Next-to-leading order QCD corrections to $bg \to tW^-$ at CERN large hadron collider," *Phys. Lett.* **B524** (2002) 283–288, `arXiv:hep-ph/0109269 [hep-ph]`. [Erratum: Phys. Lett.B537,351(2002)].

[128] J. M. Campbell and F. Tramontano, "Next-to-leading order corrections to Wt production and decay," *Nucl. Phys.* **B726** (2005) 109–130, `arXiv:hep-ph/0506289 [hep-ph]`.

[129] S. Frixione, E. Laenen, P. Motylinski, B. R. Webber, and C. D. White, "Single-top hadroproduction in association with a W boson," *JHEP* **07** (2008) 029, `arXiv:0805.3067 [hep-ph]`.

[130] Q.-H. Cao, "Demonstration of One Cutoff Phase Space Slicing Method: Next-to-Leading Order QCD Corrections to the tW Associated Production in Hadron Collision," `arXiv:0801.1539 [hep-ph]`.

[131] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method," *Eur. Phys. J.* **C71** (2011) 1547, `arXiv:1009.2450 [hep-ph]`.

[132] G. Källén, *Elementary particle physics.* Addison-Wesley, Reading, MA, 1964.

[133] C. D. White, S. Frixione, E. Laenen, and F. Maltoni, "Isolating Wt production at the LHC," *JHEP* **11** (2009) 074, `arXiv:0908.0631 [hep-ph]`.

[134] A. S. Belyaev, E. E. Boos, and L. V. Dudko, "Single top quark at future hadron colliders: Complete signal and background study," *Phys. Rev.* **D59** (1999) 075001, `arXiv:hep-ph/9806332 [hep-ph]`.

[135] B. P. Kersevan and I. Hinchliffe, "A Consistent prescription for the production involving massive quarks in hadron collisions," *JHEP* **09** (2006) 033, `arXiv:hep-ph/0603068 [hep-ph]`.

[136] N. Kauer and D. Zeppenfeld, "Finite width effects in top quark production at hadron colliders," *Phys. Rev.* **D65** (2002) 014021, `arXiv:hep-ph/0107181 [hep-ph]`.

[137] F. Demartin, B. Maier, F. Maltoni, K. Mawatari, and M. Zaro, "tWH associated production at the LHC," *Eur. Phys. J.* **C77** no. 1, (2017) 34, `arXiv:1607.05862 [hep-ph]`.

[138] A. Lazopoulos, T. McElmurry, K. Melnikov, and F. Petriello, "Next-to-leading order QCD corrections to $t\bar{t}Z$ production at the LHC," *Phys. Lett.* **B666** (2008) 62–65, `arXiv:0804.2220 [hep-ph]`.

[139] A. Kardos, Z. Trocsanyi, and C. Papadopoulos, "Top quark pair production in association with a Z-boson at NLO accuracy," *Phys. Rev.* **D85** (2012) 054015, `arXiv:1111.0610 [hep-ph]`.

[140] S. Dawson, C. Jackson, L. H. Orr, L. Reina, and D. Wackeroth, "Associated Higgs production with top quarks at the large hadron collider: NLO QCD corrections," *Phys. Rev.* **D68** (2003) 034022, `arXiv:hep-ph/0305087 [hep-ph]`.

[141] W. Beenakker, S. Dittmaier, M. Kramer, B. Plumper, M. Spira, and P. M. Zerwas, "NLO QCD corrections to t anti-t H production in hadron collisions," *Nucl. Phys.* **B653** (2003) 151–203, `arXiv:hep-ph/0211352 [hep-ph]`.

[142] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for a Standard Model Higgs Boson Produced in Association with a Top-Quark Pair and Decaying to Bottom Quarks Using a Matrix Element Method," *Eur. Phys. J.* **C75** no. 6, (2015) 251, `arXiv:1502.02485 [hep-ex]`.

[143] **ATLAS** Collaboration, G. Aad *et al.*, "Search for the Standard Model Higgs boson decaying into $b\bar{b}$ produced in association with top quarks decaying hadronically in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *JHEP* **05** (2016) 160, `arXiv:1604.03812 [hep-ex]`.

[144] M. Jezabek and J. H. Kuhn, "QCD Corrections to Semileptonic Decays of Heavy Quarks," *Nucl. Phys.* **B314** (1989) 1–6.

[145] J. A. Aguilar-Saavedra, J. Carvalho, N. F. Castro, F. Veloso, and A. Onofre, "Probing anomalous Wtb couplings in top pair decays," *Eur. Phys. J.* **C50** (2007) 519–533, `arXiv:hep-ph/0605190 [hep-ph]`.

[146] A. Czarnecki, J. G. Korner, and J. H. Piclum, "Helicity fractions of W bosons from top quark decays at NNLO in QCD," *Phys. Rev.* **D81** (2010) 111503, `arXiv:1005.2625 [hep-ph]`.

[147] *see e.g.* , M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. 1995.
`http://www.slac.stanford.edu/spires/find/books/www?cl=QC174.45%3AP4`.

[148] S. R. Coleman and E. J. Weinberg, "Radiative Corrections as the Origin of Spontaneous Symmetry Breaking," *Phys. Rev.* **D7** (1973) 1888–1910.

[149] M. Sher, "Electroweak Higgs Potentials and Vacuum Stability," *Phys. Rept.* **179** (1989) 273–418.

[150] E. Gildener, "Gauge Symmetry Hierarchies," *Phys. Rev.* **D14** (1976) 1667.

[151] S. Dimopoulos and H. Georgi, "Softly Broken Supersymmetry and SU(5)," *Nucl. Phys.* **B193** (1981) 150–162.

[152] G. 't Hooft, "Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking," *NATO Sci. Ser. B* **59** (1980) 135–157.

[153] J. R. Espinosa, G. F. Giudice, and A. Riotto, "Cosmological implications of the Higgs mass measurement," *JCAP* **0805** (2008) 002, `arXiv:0710.2484 [hep-ph]`.

[154] S. R. Coleman, "The Fate of the False Vacuum. 1. Semiclassical Theory," *Phys. Rev.* **D15** (1977) 2929–2936. [Erratum: Phys. Rev.D16,1248(1977)].

[155] C. G. Callan, Jr. and S. R. Coleman, "The Fate of the False Vacuum. 2. First Quantum Corrections," *Phys. Rev.* **D16** (1977) 1762–1768.

[156] J. Elias-Miro, J. R. Espinosa, G. F. Giudice, G. Isidori, A. Riotto, and A. Strumia, "Higgs mass implications on the stability of the electroweak vacuum," *Phys. Lett.* **B709** (2012) 222–228, `arXiv:1112.3022 [hep-ph]`.

[157] G. Degrassi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, G. Isidori, and A. Strumia, "Higgs mass and vacuum stability in the Standard Model at NNLO," *JHEP* **08** (2012) 098, `arXiv:1205.6497 [hep-ph]`.

[158] J. R. Espinosa, G. F. Giudice, E. Morgante, A. Riotto, L. Senatore, A. Strumia, and N. Tetradis, "The cosmological Higgstory of the vacuum instability," *JHEP* **09** (2015) 174, `arXiv:1505.04825 [hep-ph]`.

[159] J. Ellis, J. R. Espinosa, G. F. Giudice, A. Hoecker, and A. Riotto, "The Probable Fate of the Standard Model," *Phys. Lett.* **B679** (2009) 369–375, `arXiv:0906.0954 [hep-ph]`.

[160] M. Carena, S. Pokorski, and C. E. M. Wagner, "On the unification of couplings in the minimal supersymmetric Standard Model," *Nucl. Phys.* **B406** (1993) 59–89, `arXiv:hep-ph/9303202 [hep-ph]`.

[161] H. Georgi, H. R. Quinn, and S. Weinberg, "Hierarchy of Interactions in Unified Gauge Theories," *Phys. Rev. Lett.* **33** (1974) 451–454.

[162] S. Dimopoulos, S. Raby, and F. Wilczek, "Supersymmetry and the Scale of Unification," *Phys. Rev.* **D24** (1981) 1681–1683.

[163] P. Langacker, "Grand Unified Theories and Proton Decay," *Phys. Rept.* **72** (1981) 185.

[164] M. Dine, "TASI lectures on the strong CP problem," in *Flavor physics for the millennium. Proceedings, Theoretical Advanced Study Institute in elementary particle physics, TASI 2000, Boulder, USA, June 4-30, 2000*, pp. 349–369. 2000. `arXiv:hep-ph/0011376 [hep-ph]`. `http://alice.cern.ch/format/showfull?sysnb=2232477`.

[165] S. Weinberg, "The U(1) Problem," *Phys. Rev.* **D11** (1975) 3583–3593.

[166] G. 't Hooft, "Computation of the Quantum Effects Due to a Four-Dimensional Pseudoparticle," *Phys. Rev.* **D14** (1976) 3432–3450. [Erratum: Phys. Rev.D18,2199(1978)].

[167] G. 't Hooft, "How Instantons Solve the U(1) Problem," *Phys. Rept.* **142** (1986) 357–387.

[168] R. D. Peccei, "The Strong CP problem and axions," *Lect. Notes Phys.* **741** (2008) 3–17, `arXiv:hep-ph/0607268 [hep-ph]`. [,3(2006)].

[169] S. Weinberg, "A New Light Boson?," *Phys. Rev. Lett.* **40** (1978) 223–226.

[170] R. D. Peccei and H. R. Quinn, "Constraints Imposed by CP Conservation in the Presence of Instantons," *Phys. Rev.* **D16** (1977) 1791–1797.

[171] F. Wilczek, "Problem of Strong p and t Invariance in the Presence of Instantons," *Phys. Rev. Lett.* **40** (1978) 279–282.

[172] R. D. Peccei and H. R. Quinn, "CP Conservation in the Presence of Instantons," *Phys. Rev. Lett.* **38** (1977) 1440–1443.

[173] A. Djouadi, "The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model," *Phys. Rept.* **459** (2008) 1–241, `arXiv:hep-ph/0503173 [hep-ph]`.

[174] A. Drozd, J. Ellis, J. Quevillon, and T. You, "Comparing EFT and Exact One-Loop Analyses of Non-Degenerate Stops," *JHEP* **06** (2015) 028, `arXiv:1504.02409 [hep-ph]`.

[175] M. Schmaltz and D. Tucker-Smith, "Little Higgs review," *Ann. Rev. Nucl. Part. Sci.* **55** (2005) 229–270, `arXiv:hep-ph/0502182 [hep-ph]`.

[176] D. E. Kaplan and M. Schmaltz, "The Little Higgs from a simple group," *JHEP* **10** (2003) 039, `arXiv:hep-ph/0302049 [hep-ph]`.

[177] M. Schmaltz, "The Simplest little Higgs," *JHEP* **08** (2004) 056, `arXiv:hep-ph/0407143 [hep-ph]`.

[178] T. Han, H. E. Logan, B. McElrath, and L.-T. Wang, "Phenomenology of the little Higgs model," *Phys. Rev.* **D67** (2003) 095004, `arXiv:hep-ph/0301040 [hep-ph]`.

[179] L. Randall and R. Sundrum, "A Large mass hierarchy from a small extra dimension," *Phys. Rev. Lett.* **83** (1999) 3370–3373, `arXiv:hep-ph/9905221 [hep-ph]`.

[180] L. Randall and R. Sundrum, "An Alternative to compactification," *Phys. Rev. Lett.* **83** (1999) 4690–4693, `arXiv:hep-th/9906064 [hep-th]`.

[181] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, "The Hierarchy problem and new dimensions at a millimeter," *Phys. Lett.* **B429** (1998) 263–272, `arXiv:hep-ph/9803315 [hep-ph]`.

[182] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, "Phenomenology, astrophysics and cosmology of theories with submillimeter dimensions and TeV scale quantum gravity," *Phys. Rev.* **D59** (1999) 086004, `arXiv:hep-ph/9807344 [hep-ph]`.

[183] K. R. Dienes, E. Dudas, and T. Gherghetta, "Extra space-time dimensions and unification," *Phys. Lett.* **B436** (1998) 55–65, `arXiv:hep-ph/9803466 [hep-ph]`.

[184] T. Appelquist, H.-C. Cheng, and B. A. Dobrescu, "Bounds on universal extra dimensions," *Phys. Rev.* **D64** (2001) 035002, `arXiv:hep-ph/0012100 [hep-ph]`.

[185] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo, "Bulk gauge fields in the Randall-Sundrum model," *Phys. Lett.* **B473** (2000) 43–49, `arXiv:hep-ph/9911262 [hep-ph]`.

[186] T. Gherghetta and A. Pomarol, "Bulk fields and supersymmetry in a slice of AdS," *Nucl. Phys.* **B586** (2000) 141–162, `arXiv:hep-ph/0003129 [hep-ph]`.

[187] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo, "Phenomenology of the Randall-Sundrum Gauge Hierarchy Model," *Phys. Rev. Lett.* **84** (2000) 2080, `arXiv:hep-ph/9909255 [hep-ph]`.

[188] A. Pomarol, "Gauge bosons in a five-dimensional theory with localized gravity," *Phys. Lett.* **B486** (2000) 153–157, `arXiv:hep-ph/9911294 [hep-ph]`.

[189] A. V. Manohar, "Effective field theories," in *Quarks and colliders. Proceedings, 10th Winter Institute, Lake Louise, Canada, February 19-25, 1995*, pp. 274–315. 1995. `arXiv:hep-ph/9508245 [hep-ph]`.

[190] C. P. Burgess, "Introduction to Effective Field Theory," *Ann. Rev. Nucl. Part. Sci.* **57** (2007) 329–362, `arXiv:hep-th/0701053 [hep-th]`.

[191] D. B. Kaplan, "Effective field theories," in *Beyond the standard model 5. Proceedings, 5th Conference, Balholm, Norway, April 29-May 4, 1997*. 1995. `arXiv:nucl-th/9506035 [nucl-th]`.

[192] K. G. Wilson, "Renormalization group and critical phenomena. 1. Renormalization group and the Kadanoff scaling picture," *Phys. Rev.* **B4** (1971) 3174–3183.

[193] K. G. Wilson, "Renormalization group and critical phenomena. 2. Phase space cell analysis of critical behavior," *Phys. Rev.* **B4** (1971) 3184–3205.

[194] K. G. Wilson and J. B. Kogut, "The Renormalization group and the epsilon expansion," *Phys. Rept.* **12** (1974) 75–200.

[195] C. J. C. Burges and H. J. Schnitzer, "Virtual Effects of Excited Quarks as Probes of a Possible New Hadronic Mass Scale," *Nucl. Phys.* **B228** (1983) 464–500.

[196] C. N. Leung, S. T. Love, and S. Rao, "Low-Energy Manifestations of a New Interaction Scale: Operator Analysis," *Z. Phys.* **C31** (1986) 433.

[197] W. Buchmuller and D. Wyler, "Effective Lagrangian Analysis of New Interactions and Flavor Conservation," *Nucl. Phys.* **B268** (1986) 621–653.

[198] S. Weinberg, "Baryon and Lepton Nonconserving Processes," *Phys. Rev. Lett.* **43** (1979) 1566–1570.

[199] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek, "Dimension-Six Terms in the Standard Model Lagrangian," *JHEP* **10** (2010) 085, `arXiv:1008.4884 [hep-ph]`.

[200] C. Grosse-Knetter, "Effective Lagrangians with higher derivatives and equations of motion," *Phys. Rev.* **D49** (1994) 6709–6719, `arXiv:hep-ph/9306321 [hep-ph]`.

[201] H. Simma, "Equations of motion for effective Lagrangians and penguins in rare B decays," *Z. Phys.* **C61** (1994) 67–82, `arXiv:hep-ph/9307274 [hep-ph]`.

[202] J. Wudka, "Electroweak effective Lagrangians," *Int. J. Mod. Phys.* **A9** (1994) 2301–2362, `arXiv:hep-ph/9406205 [hep-ph]`.

[203] C. Arzt, M. B. Einhorn, and J. Wudka, "Patterns of deviation from the standard model," *Nucl. Phys.* **B433** (1995) 41–66, `arXiv:hep-ph/9405214 [hep-ph]`.

[204] J. A. Aguilar-Saavedra, "A Minimal set of top anomalous couplings," *Nucl. Phys.* **B812** (2009) 181–204, `arXiv:0811.3842 [hep-ph]`.

[205] J. A. Aguilar-Saavedra, "A Minimal set of top-Higgs anomalous couplings," *Nucl. Phys.* **B821** (2009) 215–227, `arXiv:0904.2387 [hep-ph]`.

[206] J. A. Aguilar-Saavedra, "Effective four-fermion operators in top physics: A Roadmap," *Nucl. Phys.* **B843** (2011) 638–672, `arXiv:1008.3562 [hep-ph]`. [Erratum: Nucl. Phys.B851,443(2011)].

[207] P. J. Fox, Z. Ligeti, M. Papucci, G. Perez, and M. D. Schwartz, "Deciphering top flavor violation at the LHC with B factories," *Phys. Rev.* **D78** (2008) 054008, `arXiv:0704.1482 [hep-ph]`.

[208] K. Hagiwara, S. Ishihara, R. Szalapski, and D. Zeppenfeld, "Low-energy effects of new interactions in the electroweak boson sector," *Phys. Rev.* **D48** (1993) 2182–2203.

[209] G. F. Giudice, C. Grojean, A. Pomarol, and R. Rattazzi, "The Strongly-Interacting Light Higgs," *JHEP* **06** (2007) 045, `arXiv:hep-ph/0703164 [hep-ph]`.

[210] P. Artoisenet *et al.*, "A framework for Higgs characterisation," *JHEP* **11** (2013) 043, `arXiv:1306.6464 [hep-ph]`.

[211] R. S. Gupta, A. Pomarol, and F. Riva, "BSM Primary Effects," *Phys. Rev.* **D91** no. 3, (2015) 035001, `arXiv:1405.0181 [hep-ph]`.

[212] A. Falkowski, B. Fuks, K. Mawatari, K. Mimasu, F. Riva, and V. sanz, "Rosetta: an operator basis translator for Standard Model effective field theory," *Eur. Phys. J.* **C75** no. 12, (2015) 583, `arXiv:1508.05895 [hep-ph]`.

[213] B. Henning, X. Lu, T. Melia, and H. Murayama, "2, 84, 30, 993, 560, 15456, 11962, 261485, ...: Higher dimension operators in the SM EFT," `arXiv:1512.03433 [hep-ph]`.

[214] C. Degrande, J.-M. Gerard, C. Grojean, F. Maltoni, and G. Servant, "Non-resonant New Physics in Top Pair Production at Hadron Colliders," *JHEP* **03** (2011) 125, `arXiv:1010.6304 [hep-ph]`.

[215] P. L. Cho and E. H. Simmons, "Looking for gluon substructure at the tevatron," *Phys. Lett.* **B323** (1994) 401–407, `arXiv:hep-ph/9307345 [hep-ph]`.

[216] P. L. Cho and E. H. Simmons, "Searching for G3 in $t\bar{t}$ production," *Phys. Rev.* **D51** (1995) 2360–2370, `arXiv:hep-ph/9408206 [hep-ph]`.

[217] C. Zhang and S. Willenbrock, "Effective-Field-Theory Approach to Top-Quark Production and Decay," *Phys. Rev.* **D83** (2011) 034006, `arXiv:1008.3869 [hep-ph]`.

[218] T. Appelquist and J. Carazzone, "Infrared Singularities and Massive Fields," *Phys. Rev.* **D11** (1975) 2856.

[219] K. G. Wilson, "The renormalization group and critical phenomena," *Rev. Mod. Phys.* **55** (1983) 583–600.

[220] K. Hagiwara, R. Peccei, D. Zeppenfeld, and K. Hikasa, "Probing the Weak Boson Sector in $e^+e^- \to W^+W^-$," *Nucl.Phys.* **B282** (1987) 253.

[221] S. F. King, M. Muhlleitner, R. Nevzorov, and K. Walz, "Natural NMSSM Higgs Bosons," *Nucl. Phys.* **B870** (2013) 323–352, `arXiv:1211.5074 [hep-ph]`.

[222] A. Azatov, R. Contino, and J. Galloway, "Model-Independent Bounds on a Light Higgs," *JHEP* **04** (2012) 127, `arXiv:1202.3415 [hep-ph]`. [Erratum: JHEP04,140(2013)].

[223] J. R. Espinosa, C. Grojean, M. Muhlleitner, and M. Trott, "First Glimpses at Higgs' face," *JHEP* **12** (2012) 045, `arXiv:1207.1717 [hep-ph]`.

[224] T. Plehn and M. Rauch, "Higgs Couplings after the Discovery," *Europhys. Lett.* **100** (2012) 11002, `arXiv:1207.6108 [hep-ph]`.

[225] D. Carmi, A. Falkowski, E. Kuflik, T. Volansky, and J. Zupan, "Higgs After the Discovery: A Status Report," *JHEP* **10** (2012) 196, `arXiv:1207.1718 [hep-ph]`.

[226] M. E. Peskin, "Comparison of LHC and ILC Capabilities for Higgs Boson Coupling Measurements," `arXiv:1207.2516 [hep-ph]`.

[227] B. Dumont, S. Fichet, and G. von Gersdorff, "A Bayesian view of the Higgs sector with higher dimensional operators," *JHEP* **07** (2013) 065, `arXiv:1304.3369 [hep-ph]`.

[228] A. Djouadi and G. Moreau, "The couplings of the Higgs boson and its CP properties from fits of the signal strengths and their ratios at the 7+8 TeV LHC," *Eur. Phys. J.* **C73** no. 9, (2013) 2512, `arXiv:1303.6591 [hep-ph]`.

[229] D. Lopez-Val, T. Plehn, and M. Rauch, "Measuring extended Higgs sectors as a consistent free couplings model," *JHEP* **10** (2013) 134, `arXiv:1308.1979 [hep-ph]`.

[230] C. Englert, A. Freitas, M. M. Muehlleitner, T. Plehn, M. Rauch, M. Spira, and K. Walz, "Precision Measurements of Higgs Couplings: Implications for New Physics Scales," *J. Phys.* **G41** (2014) 113001, `arXiv:1403.7191 [hep-ph]`.

[231] J. Ellis, V. Sanz, and T. You, "Complete Higgs Sector Constraints on Dimension-6 Operators," *JHEP* **07** (2014) 036, `arXiv:1404.3667 [hep-ph]`.

[232] J. Ellis, V. Sanz, and T. You, "The Effective Standard Model after LHC Run I," *JHEP* **03** (2015) 157, `arXiv:1410.7703 [hep-ph]`.

[233] A. Falkowski and F. Riva, "Model-independent precision constraints on dimension-6 operators," *JHEP* **02** (2015) 039, `arXiv:1411.0669 [hep-ph]`.

[234] T. Corbett, O. J. P. Eboli, D. Goncalves, J. Gonzalez-Fraile, T. Plehn, and M. Rauch, "The Higgs Legacy of the LHC Run I," *JHEP* **08** (2015) 156, `arXiv:1505.05516 [hep-ph]`.

[235] G. Buchalla, O. Cata, A. Celis, and C. Krause, "Fitting Higgs Data with Nonlinear Effective Theory," arXiv:1511.00988 [hep-ph].

[236] **ATLAS** Collaboration, G. Aad *et al.*, "Constraints on non-Standard Model Higgs boson interactions in an effective field theory using differential cross sections measured in the $H \to \gamma\gamma$ decay channel at $\sqrt{s} = 8$ TeV with the ATLAS detector," arXiv:1508.02507 [hep-ex].

[237] L. Berthier and M. Trott, "Consistent constraints on the Standard Model Effective Field Theory," arXiv:1508.05060 [hep-ph].

[238] A. Falkowski, M. Gonzalez-Alonso, A. Greljo, and D. Marzocca, "Global constraints on anomalous triple gauge couplings in effective field theory approach," arXiv:1508.00581 [hep-ph].

[239] C. Englert, R. Kogler, H. Schulz, and M. Spannowsky, "Higgs coupling measurements at the LHC," arXiv:1511.05170 [hep-ph].

[240] C. Bernardo, N. F. Castro, M. C. N. Fiolhais, H. Gonçalves, A. G. C. Guerra, *et al.*, "Studying the Wtb vertex structure using recent LHC results," *Phys.Rev.* **D90** no. 11, (2014) 113007, arXiv:1408.7063 [hep-ph].

[241] B. Grzadkowski, Z. Hioki, K. Ohkuma, and J. Wudka, "Probing anomalous top quark couplings induced by dimension-six operators at photon colliders," *Nucl. Phys.* **B689** (2004) 108–126, arXiv:hep-ph/0310159 [hep-ph].

[242] D. Nomura, "Effects of Top-quark Compositeness on Higgs Boson Production at the LHC," *JHEP* **02** (2010) 061, arXiv:0911.1941 [hep-ph].

[243] Z. Hioki and K. Ohkuma, "Search for anomalous top-gluon couplings at LHC revisited," *Eur. Phys. J.* **C65** (2010) 127–135, arXiv:0910.3049 [hep-ph].

[244] Z. Hioki and K. Ohkuma, "Addendum to: Search for anomalous top-gluon couplings at LHC revisited," *Eur. Phys. J.* **C71** (2011) 1535, arXiv:1011.2655 [hep-ph].

[245] Z. Hioki and K. Ohkuma, "Latest constraint on nonstandard top-gluon couplings at hadron colliders and its future prospect," *Phys. Rev.* **D88** (2013) 017503, arXiv:1306.5387 [hep-ph].

[246] J. A. Aguilar-Saavedra, B. Fuks, and M. L. Mangano, "Pinning down top dipole moments with ultra-boosted tops," *Phys. Rev.* **D91** (2015) 094021, arXiv:1412.6654 [hep-ph].

[247] C.-R. Chen, F. Larios, and C. P. Yuan, "General analysis of single top production and W helicity in top decay," *Phys. Lett.* **B631** (2005) 126–132, `arXiv:hep-ph/0503040 [hep-ph]`.

[248] J. A. Aguilar-Saavedra, "Single top quark production at LHC with anomalous Wtb couplings," *Nucl. Phys.* **B804** (2008) 160–192, `arXiv:0803.3810 [hep-ph]`.

[249] J. A. Aguilar-Saavedra and J. Bernabeu, "W polarisation beyond helicity fractions in top quark decays," *Nucl. Phys.* **B840** (2010) 349–378, `arXiv:1005.5382 [hep-ph]`.

[250] J. A. Aguilar-Saavedra, N. F. Castro, and A. Onofre, "Constraints on the Wtb vertex from early LHC data," *Phys. Rev.* **D83** (2011) 117301, `arXiv:1105.0117 [hep-ph]`.

[251] M. Fabbrichesi, M. Pinamonti, and A. Tonero, "Limits on anomalous top quark gauge couplings from Tevatron and LHC data," *Eur. Phys. J.* **C74** no. 12, (2014) 3193, `arXiv:1406.5393 [hep-ph]`.

[252] M. Fabbrichesi, M. Pinamonti, and A. Tonero, "Stringent limits on top-quark compositeness from $t\bar{t}$ production at the Tevatron and the LHC," *Phys. Rev.* **D89** no. 7, (2014) 074028, `arXiv:1307.5750 [hep-ph]`.

[253] Q.-H. Cao, B. Yan, J.-H. Yu, and C. Zhang, "A General Analysis of Wtb anomalous Couplings," `arXiv:1504.03785 [hep-ph]`.

[254] G. A. Gonzalez-Sprinberg, R. Martinez, and J. Vidal, "Top quark tensor couplings," *JHEP* **07** (2011) 094, `arXiv:1105.5601 [hep-ph]`. [Erratum: JHEP05,117(2013)].

[255] S. Davidson, M. L. Mangano, S. Perries, and V. Sordini, "Lepton Flavour Violating top decays at the LHC," *Eur. Phys. J.* **C75** no. 9, (2015) 450, `arXiv:1507.07163 [hep-ph]`.

[256] S. Jung, P. Ko, Y. W. Yoon, and C. Yu, "Renormalization group-induced phenomena of top pairs from four-quark effective operators," *JHEP* **08** (2014) 120, `arXiv:1406.4570 [hep-ph]`.

[257] Q.-H. Cao, J. Wudka, and C. P. Yuan, "Search for new physics via single top production at the LHC," *Phys. Lett.* **B658** (2007) 50–56, `arXiv:0704.2809 [hep-ph]`.

[258] N. Greiner, S. Willenbrock, and C. Zhang, "Effective Field Theory for Nonstandard Top Quark Couplings," *Phys. Lett.* **B704** (2011) 218–222, `arXiv:1104.3122 [hep-ph]`.

[259] C. Degrande, N. Greiner, W. Kilian, O. Mattelaer, H. Mebane, T. Stelzer, S. Willenbrock, and C. Zhang, "Effective Field Theory: A Modern Approach to Anomalous Couplings," *Annals Phys.* **335** (2013) 21–32, `arXiv:1205.4231 [hep-ph]`.

[260] J. de Blas, M. Chala, and J. Santiago, "Renormalization Group Constraints on New Top Interactions from Electroweak Precision Data," *JHEP* **09** (2015) 189, `arXiv:1507.00757 [hep-ph]`.

[261] R. R. Aguilar, A. O. Bouzas, and F. Larios, "Limits on the anomalous Wtq couplings," `arXiv:1509.06431 [hep-ph]`.

[262] G. Durieux, F. Maltoni, and C. Zhang, "Global approach to top-quark flavor-changing interactions," *Phys. Rev.* **D91** no. 7, (2015) 074017, `arXiv:1412.7166 [hep-ph]`.

[263] R. Contino, M. Ghezzi, C. Grojean, M. Muhlleitner, and M. Spira, "Effective Lagrangian for a light Higgs-like scalar," *JHEP* **07** (2013) 035, `arXiv:1303.3876 [hep-ph]`.

[264] E. Masso, "An Effective Guide to Beyond the Standard Model Physics," *JHEP* **10** (2014) 128, `arXiv:1406.6376 [hep-ph]`.

[265] A. Pomarol, "Higgs Physics," in *2014 European School of High-Energy Physics (ESHEP 2014) Garderen, The Netherlands, June 18-July 1, 2014*. 2014. `arXiv:1412.4410 [hep-ph]`.

[266] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurements of $t\bar{t}$ spin correlations and top-quark polarization using dilepton final states in pp collisions at $\sqrt{s} = 7$ TeV," *Phys. Rev. Lett.* **112** no. 18, (2014) 182001, `arXiv:1311.3924 [hep-ex]`.

[267] **ATLAS** Collaboration, G. Aad *et al.*, "Measurements of spin correlation in top-antitop quark events from proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector," *Phys. Rev.* **D90** no. 11, (2014) 112016, `arXiv:1407.4314 [hep-ex]`.

[268] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of Spin Correlation in Top-Antitop Quark Events and Search for Top Squark Pair Production in pp Collisions at $\sqrt{s} = 8$ TeV Using the ATLAS Detector," *Phys. Rev. Lett.* **114** no. 14, (2015) 142001, `arXiv:1412.4742 [hep-ex]`.

[269] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b -tagged jets in pp collisions at $\sqrt{s} = 7$ and 8

TeV with the ATLAS detector," *Eur. Phys. J.* **C74** no. 10, (2014) 3109,
`arXiv:1406.5375 [hep-ex]`.

[270] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of the cross section for
top-quark pair production in *pp* collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector
using final states with two high-pt leptons," *JHEP* **05** (2012) 059,
`arXiv:1202.4892 [hep-ex]`.

[271] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of the top quark pair cross
section with ATLAS in *pp* collisions at $\sqrt{s} = 7$ TeV using final states with an
electron or a muon and a hadronically decaying τ lepton," *Phys. Lett.* **B717** (2012)
89–108, `arXiv:1205.2067 [hep-ex]`.

[272] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of the top quark pair
production cross-section with ATLAS in the single lepton channel," *Phys. Lett.*
B711 (2012) 244–263, `arXiv:1201.1889 [hep-ex]`.

[273] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of the ttbar production cross
section in the tau+jets channel using the ATLAS detector," *Eur. Phys. J.* **C73**
no. 3, (2013) 2328, `arXiv:1211.7205 [hep-ex]`.

[274] **ATLAS** Collaboration, G. Aad *et al.*, "Simultaneous measurements of the $t\bar{t}$,
W^+W^-, and $Z/\gamma^* \to \tau\tau$ production cross-sections in *pp* collisions at $\sqrt{s} = 7$ TeV
with the ATLAS detector," *Phys. Rev.* **D91** no. 5, (2015) 052005,
`arXiv:1407.0573 [hep-ex]`.

[275] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of the top pair production
cross section in 8 TeV proton-proton collisions using kinematic information in the
lepton+jets final state with ATLAS," *Phys. Rev.* **D91** no. 11, (2015) 112013,
`arXiv:1504.04251 [hep-ex]`.

[276] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of the $t\bar{t}$ production cross
section in the all-jet final state in pp collisions at $\sqrt{s} = 7$ TeV," *JHEP* **05** (2013)
065, `arXiv:1302.0508 [hep-ex]`.

[277] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of the $t\bar{t}$ production cross
section in the dilepton channel in *pp* collisions at $\sqrt{s} = 7$ TeV," *JHEP* **11** (2012)
067, `arXiv:1208.2671 [hep-ex]`.

[278] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of the $t\bar{t}$ production cross
section in *pp* collisions at $\sqrt{s} = 7$ TeV with lepton + jets final states," *Phys. Lett.*
B720 (2013) 83–104, `arXiv:1212.6682`.

[279] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of the top quark pair production cross section in *pp* collisions at $\sqrt{s} = 7$ TeV in dilepton final states containing a τ," *Phys. Rev.* **D85** (2012) 112007, arXiv:1203.6810 [hep-ex].

[280] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of the top-antitop production cross section in the tau+jets channel in pp collisions at sqrt(s) = 7 TeV," *Eur. Phys. J.* **C73** no. 4, (2013) 2386, arXiv:1301.5755 [hep-ex].

[281] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 8$ TeV," *JHEP* **02** (2014) 024, arXiv:1312.7582 [hep-ex]. [Erratum: JHEP02,102(2014)].

[282] **CMS** Collaboration, V. Khachatryan *et al.*, "Measurement of the top quark pair production cross section in proton-proton collisions at $\sqrt{s} = 13$ TeV," arXiv:1510.05302 [hep-ex].

[283] **CDF, D0** Collaboration, T. A. Aaltonen *et al.*, "Combination of measurements of the top-quark pair production cross section from the Tevatron Collider," *Phys. Rev.* **D89** no. 7, (2014) 072001, arXiv:1309.7570 [hep-ex].

[284] **ATLAS** Collaboration, G. Aad *et al.*, "Comprehensive measurements of t-channel single top-quark production cross sections at $\sqrt{s} = 7$ TeV with the ATLAS detector," *Phys. Rev.* **D90** no. 11, (2014) 112006, arXiv:1406.7844 [hep-ex].

[285] **CDF** Collaboration, T. A. Aaltonen *et al.*, "Evidence for s-channel Single-Top-Quark Production in Events with one Charged Lepton and two Jets at CDF," *Phys. Rev. Lett.* **112** (2014) 231804, arXiv:1402.0484 [hep-ex].

[286] **CMS** Collaboration, V. Khachatryan *et al.*, "Measurement of the t-channel single-top-quark production cross section and of the $|V_{tb}|$ CKM matrix element in pp collisions at $\sqrt{s}= 8$ TeV," *JHEP* **06** (2014) 090, arXiv:1403.7366 [hep-ex].

[287] **D0** Collaboration, V. M. Abazov *et al.*, "Measurement of the t-channel single top quark production cross section," *Phys. Lett.* **B682** (2010) 363–369, arXiv:0907.4259 [hep-ex].

[288] **D0** Collaboration, V. M. Abazov *et al.*, "Model-independent measurement of t-channel single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," *Phys. Lett.* **B705** (2011) 313–319, arXiv:1105.2788 [hep-ex].

[289] **ATLAS** Collaboration, G. Aad *et al.*, "Measurements of normalized differential cross sections for $t\bar{t}$ production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector," *Phys. Rev.* **D90** no. 7, (2014) 072004, arXiv:1407.0371 [hep-ex].

[290] **CDF** Collaboration, T. Aaltonen *et al.*, "First Measurement of the t anti-t Differential Cross Section d sigma/dM(t anti-t) in p anti-p Collisions at s**(1/2)=1.96-TeV," *Phys. Rev. Lett.* **102** (2009) 222003, arXiv:0903.2850 [hep-ex].

[291] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of differential top-quark pair production cross sections in pp colisions at $\sqrt{s} = 7$ TeV," *Eur. Phys. J.* **C73** no. 3, (2013) 2339, arXiv:1211.2220 [hep-ex].

[292] **CMS** Collaboration, V. Khachatryan *et al.*, "Measurement of the Differential Cross Section for Top Quark Pair Production in pp Collisions at $\sqrt{s} = 8$ TeV," arXiv:1505.04480 [hep-ex].

[293] **D0** Collaboration, V. M. Abazov *et al.*, "Measurement of differential $t\bar{t}$ production cross sections in $p\bar{p}$ collisions," *Phys. Rev.* **D90** no. 9, (2014) 092006, arXiv:1401.5785 [hep-ex].

[294] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of the top quark pair production charge asymmetry in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector," *JHEP* **02** (2014) 107, arXiv:1311.6724 [hep-ex].

[295] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurements of the $t\bar{t}$ charge asymmetry using the dilepton decay channel in pp collisions at $\sqrt{s} = 7$ TeV," *JHEP* **04** (2014) 191, arXiv:1402.3803 [hep-ex].

[296] **CDF** Collaboration, T. Aaltonen *et al.*, "Measurement of the top quark forward-backward production asymmetry and its dependence on event kinematic properties," *Phys. Rev.* **D87** no. 9, (2013) 092002, arXiv:1211.1003 [hep-ex].

[297] **D0** Collaboration, V. M. Abazov *et al.*, "Measurement of the forward-backward asymmetry in top quark-antiquark production in ppbar collisions using the lepton+jets channel," *Phys. Rev.* **D90** (2014) 072011, arXiv:1405.0421 [hep-ex].

[298] **CDF** Collaboration, T. A. Aaltonen *et al.*, "Direct Measurement of the Total Decay Width of the Top Quark," *Phys. Rev. Lett.* **111** no. 20, (2013) 202001, arXiv:1308.4050 [hep-ex].

[299] **D0** Collaboration, V. M. Abazov *et al.*, "An Improved determination of the width of the top quark," *Phys. Rev.* **D85** (2012) 091104, arXiv:1201.4156 [hep-ex].

[300] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of the W boson polarization in top quark decays with the ATLAS detector," *JHEP* **06** (2012) 088, arXiv:1205.2484 [hep-ex].

[301] **CDF** Collaboration, T. Aaltonen *et al.*, "Measurement of *W*-Boson Polarization in Top-quark Decay using the Full CDF Run II Data Set," *Phys. Rev.* **D87** no. 3, (2013) 031104, `arXiv:1211.4523 [hep-ex]`.

[302] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of the W-boson helicity in top-quark decays from $t\bar{t}$ production in lepton+jets events in pp collisions at $\sqrt{s} = 7$ TeV," *JHEP* **10** (2013) 167, `arXiv:1308.3879 [hep-ex]`.

[303] **D0** Collaboration, V. M. Abazov *et al.*, "Measurement of the W boson helicity in top quark decays using 5.4 fb^{-1} of $p\bar{p}$ collision data," *Phys. Rev.* **D83** (2011) 032009, `arXiv:1011.6549 [hep-ex]`.

[304] **ATLAS** Collaboration, G. Aad *et al.*, "Observation of top-quark pair production in association with a photon and measurement of the $t\bar{t}\gamma$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector," *Phys. Rev.* **D91** no. 7, (2015) 072007, `arXiv:1502.00586 [hep-ex]`.

[305] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of the $t\bar{t}W$ and $t\bar{t}Z$ production cross sections in *pp* collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," `arXiv:1509.05276 [hep-ex]`.

[306] **CMS** Collaboration, V. Khachatryan *et al.*, "Measurement of top quark-antiquark pair production in association with a W or Z boson in pp collisions at $\sqrt{s} = 8$ TeV," *Eur. Phys. J.* **C74** no. 9, (2014) 3060, `arXiv:1406.7830 [hep-ex]`.

[307] A. Buckley, H. Hoeth, H. Lacker, H. Schulz, and J. E. von Seggern, "Systematic event generator tuning for the LHC," *Eur.Phys.J.* **C65** (2010) 331–357, `arXiv:0907.2973 [hep-ph]`.

[308] J. Butterworth *et al.*, "PDF4LHC recommendations for LHC Run II," `arXiv:1510.03865 [hep-ph]`.

[309] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, *et al.*, "Implications of CTEQ global analysis for collider observables," *Phys.Rev.* **D78** (2008) 013004, `arXiv:0802.0007 [hep-ph]`.

[310] R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, *et al.*, "A first unbiased global NLO determination of parton distributions and their uncertainties," *Nucl.Phys.* **B838** (2010) 136–206, `arXiv:1002.4407 [hep-ph]`.

[311] G. Passarino, "NLO Inspired Effective Lagrangians for Higgs Physics," *Nucl. Phys.* (2013) 416–458, `arXiv:1209.5538 [hep-ph]`.

[312] H. Mebane, N. Greiner, C. Zhang, and S. Willenbrock, "Constraints on Electroweak Effective Operators at One Loop," *Phys.Rev.* **D88** no. 1, (2013) 015028, `1306.3380`.

[313] E. E. Jenkins, A. V. Manohar, and M. Trott, "Renormalization Group Evolution of the Standard Model Dimension Six Operators I: Formalism and lambda Dependence," *JHEP* **10** (2013) 087, `arXiv:1308.2627 [hep-ph]`.

[314] E. E. Jenkins, A. V. Manohar, and M. Trott, "Naive Dimensional Analysis Counting of Gauge Theory Amplitudes and Anomalous Dimensions," *Phys. Lett.* **B726** (2013) 697–702, `arXiv:1309.0819 [hep-ph]`.

[315] E. E. Jenkins, A. V. Manohar, and M. Trott, "Renormalization Group Evolution of the Standard Model Dimension Six Operators II: Yukawa Dependence," *JHEP* **01** (2014) 035, `arXiv:1310.4838 [hep-ph]`.

[316] R. Alonso, E. E. Jenkins, A. V. Manohar, and M. Trott, "Renormalization Group Evolution of the Standard Model Dimension Six Operators III: Gauge Coupling Dependence and Phenomenology," *JHEP* **04** (2014) 159, `arXiv:1312.2014 [hep-ph]`.

[317] C. Hartmann and M. Trott, "On one-loop corrections in the standard model effective field theory; the $\Gamma(h \to \gamma\,\gamma)$ case," *JHEP* **07** (2015) 151, `arXiv:1505.02646 [hep-ph]`.

[318] M. Ghezzi, R. Gomez-Ambrosio, G. Passarino, and S. Uccirati, "NLO Higgs effective field theory and kappa-framework," *JHEP* **07** (2015) 175, `arXiv:1505.03706 [hep-ph]`.

[319] C. Zhang and F. Maltoni, "Top-quark decay into Higgs boson and a light quark at next-to-leading order in QCD," *Phys. Rev.* **D88** (2013) 054005, `arXiv:1305.7386 [hep-ph]`.

[320] C. Englert and M. Spannowsky, "Effective Theories and Measurements at Colliders," *Phys. Lett.* **B740** (2015) 8–15, `arXiv:1408.5147 [hep-ph]`.

[321] C. Hartmann and M. Trott, "Higgs decay to two photons at one-loop in the SMEFT," `arXiv:1507.03568 [hep-ph]`.

[322] C. Cheung and C.-H. Shen, "Nonrenormalization Theorems without Supersymmetry," *Phys. Rev. Lett.* **115** no. 7, (2015) 071601, `arXiv:1505.01844 [hep-ph]`.

[323] A. Drozd, J. Ellis, J. Quevillon, and T. You, "The Universal One-Loop Effective Action," `arXiv:1512.03003 [hep-ph]`.

[324] R. Gauld, B. D. Pecjak, and D. J. Scott, "One-loop corrections to $h \to b\bar{b}$ and $h \to \tau\bar{\tau}$ decays in the Standard Model Dimension-6 EFT: four-fermion operators and the large-m_t limit," `arXiv:1512.02508 [hep-ph]`.

[325] L. Berthier and M. Trott, "Towards consistent Electroweak Precision Data constraints in the SMEFT," *JHEP* **05** (2015) 024, `arXiv:1502.02570 [hep-ph]`.

[326] N. D. Christensen and C. Duhr, "FeynRules - Feynman rules made easy," *Comput.Phys.Commun.* **180** (2009) 1614–1641, `arXiv:0806.4194 [hep-ph]`.

[327] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, *et al.*, "UFO - The Universal FeynRules Output," *Comput.Phys.Commun.* **183** (2012) 1201–1214, `arXiv:1108.2040 [hep-ph]`.

[328] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations," *JHEP* **06** (2002) 029, `arXiv:hep-ph/0204244 [hep-ph]`.

[329] C. Zhang and S. Willenbrock, "Effective Field Theory for Top Quark Physics," *Nuovo Cim.* **C033** no. 4, (2010) 285–291, `arXiv:1008.3155 [hep-ph]`.

[330] **CDF** Collaboration, T. Aaltonen *et al.*, "Evidence for a Mass Dependent Forward-Backward Asymmetry in Top Quark Pair Production," *Phys. Rev.* **D83** (2011) 112003, `arXiv:1101.0034 [hep-ex]`.

[331] W. Hollik and D. Pagani, "The electroweak contribution to the top quark forward-backward asymmetry at the Tevatron," *Phys. Rev.* **D84** (2011) 093003, `arXiv:1107.2606 [hep-ph]`.

[332] M. Bauer, F. Goertz, U. Haisch, T. Pfoh, and S. Westhoff, "Top-Quark Forward-Backward Asymmetry in Randall-Sundrum Models Beyond the Leading Order," *JHEP* **11** (2010) 039, `arXiv:1008.0742 [hep-ph]`.

[333] C. Delaunay, O. Gedalia, Y. Hochberg, G. Perez, and Y. Soreq, "Implications of the CDF $t\bar{t}$ Forward-Backward Asymmetry for Hard Top Physics," *JHEP* **08** (2011) 031, `arXiv:1103.2297 [hep-ph]`.

[334] J. A. Aguilar-Saavedra, A. Juste, and F. Rubbo, "Boosting the $t\bar{t}$ charge asymmetry," *Phys. Lett.* **B707** (2012) 92–98, `arXiv:1109.3710 [hep-ph]`.

[335] A. Biekötter, J. Brehmer, and T. Plehn, "Extending the limits of Higgs effective theory," *Phys. Rev.* **D94** no. 5, (2016) 055032, `arXiv:1602.05202 [hep-ph]`.

[336] A. Freitas, D. López-Val, and T. Plehn, "When matching matters: Loop effects in Higgs effective theory," *Phys. Rev.* **D94** no. 9, (2016) 095007, `arXiv:1607.08251 [hep-ph]`.

[337] J. Brehmer, A. Freitas, D. Lopez-Val, and T. Plehn, "Pushing Higgs Effective Theory to its Limits," `arXiv:1510.03443 [hep-ph]`.

[338] G. Isidori and M. Trott, "Higgs form factors in Associated Production," *JHEP* **02** (2014) 082, `arXiv:1307.4051 [hep-ph]`.

[339] E. Boos, V. Bunichev, L. Dudko, and M. Perfilov, "Interference between W' and W in single-top quark production processes," *Phys. Lett.* **B655** (2007) 245–250, `arXiv:hep-ph/0610080 [hep-ph]`.

[340] **ATLAS** Collaboration, G. Aad *et al.*, "Search for New Physics in the Dijet Mass Distribution using 1 fb^{-1} of pp Collision Data at $\sqrt{s}$ =7 TeV collected by the ATLAS Detector," *Phys. Lett.* **B708** (2012) 37–54, `arXiv:1108.6311 [hep-ex]`.

[341] **ATLAS** Collaboration, G. Aad *et al.*, "Search for new phenomena in the dijet mass distribution using $p - p$ collision data at $\sqrt{s} = 8$ TeV with the ATLAS detector," *Phys. Rev.* **D91** no. 5, (2015) 052007, `arXiv:1407.1376 [hep-ex]`.

[342] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for resonances and quantum black holes using dijet mass spectra in proton-proton collisions at $\sqrt{s} = 8$ TeV," *Phys. Rev.* **D91** no. 5, (2015) 052009, `arXiv:1501.04198 [hep-ex]`.

[343] T. Plehn, G. P. Salam, and M. Spannowsky, "Fat Jets for a Light Higgs," *Phys. Rev. Lett.* **104** (2010) 111801, `arXiv:0910.5472 [hep-ph]`.

[344] T. Plehn, M. Spannowsky, M. Takeuchi, and D. Zerwas, "Stop Reconstruction with Tagged Tops," *JHEP* **10** (2010) 078, `arXiv:1006.2833 [hep-ph]`.

[345] T. Plehn and M. Spannowsky, "Top Tagging," *J. Phys.* **G39** (2012) 083001, `arXiv:1112.4441 [hep-ph]`.

[346] A. Altheimer *et al.*, "Boosted objects and jet substructure at the LHC. Report of BOOST2012, held at IFIC Valencia, 23rd-27th of July 2012," *Eur. Phys. J.* **C74** no. 3, (2014) 2792, `arXiv:1311.2708 [hep-ex]`.

[347] S. Schaetzel and M. Spannowsky, "Tagging highly boosted top quarks," *Phys. Rev.* **D89** no. 1, (2014) 014007, `arXiv:1308.0540 [hep-ph]`.

[348] T. Plehn, M. Spannowsky, and M. Takeuchi, "How to Improve Top Tagging," *Phys. Rev.* **D85** (2012) 034029, `arXiv:1111.5034 [hep-ph]`.

[349] M. Backovic, O. Gabizon, J. Juknevich, G. Perez, and Y. Soreq, "Measuring boosted tops in semi-leptonic $t\bar{t}$ events for the standard model and beyond," *JHEP* **04** (2014) 176, `arXiv:1311.2962 [hep-ph]`.

[350] M. Bahr, S. Gieseke, M. Gigg, D. Grellscheid, K. Hamilton, *et al.*, "Herwig++ Physics and Manual," *Eur.Phys.J.* **C58** (2008) 639–707, `arXiv:0803.0883 [hep-ph]`.

[351] J. Bellm *et al.*, "Herwig 7.0/Herwig++ 3.0 release note," *Eur. Phys. J.* **C76** no. 4, (2016) 196, `arXiv:1512.01178 [hep-ph]`.

[352] O. Bessidskaia Bylund, F. Maltoni, I. Tsinikos, E. Vryonidou, and C. Zhang, "Probing top quark neutral couplings in the Standard Model Effective Field Theory at NLO in QCD," *JHEP* **05** (2016) 052, `arXiv:1601.08193 [hep-ph]`.

[353] C. Degrande, F. Maltoni, J. Wang, and C. Zhang, "Automatic computations at next-to-leading order in QCD for top-quark flavor-changing neutral processes," *Phys. Rev.* **D91** (2015) 034024, `arXiv:1412.5594 [hep-ph]`.

[354] S. Moch, P. Uwer, and A. Vogt, "On top-pair hadro-production at next-to-next-to-leading order," *Phys. Lett.* **B714** (2012) 48–54, `arXiv:1203.6282 [hep-ph]`.

[355] A. Buckley, J. Butterworth, L. Lonnblad, D. Grellscheid, H. Hoeth, *et al.*, "Rivet user manual," *Comput.Phys.Commun.* **184** (2013) 2803–2819, `arXiv:1003.0694 [hep-ph]`.

[356] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm," *JHEP* **0804** (2008) 063, `arXiv:0802.1189 [hep-ph]`.

[357] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual," *Eur.Phys.J.* **C72** (2012) 1896, `arXiv:1111.6097 [hep-ph]`.

[358] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, "Better jet clustering algorithms," *JHEP* **08** (1997) 001, `arXiv:hep-ph/9707323 [hep-ph]`.

[359] M. Wobisch and T. Wengler, "Hadronization corrections to jet cross-sections in deep inelastic scattering," in *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998-1999*, pp. 270–279. 1998. `arXiv:hep-ph/9907280 [hep-ph]`. `http://inspirehep.net/record/484872/files/arXiv:hep-ph_9907280.pdf`.

[360] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector,".

[361] G. Kasieczka, T. Plehn, T. Schell, T. Strebler, and G. P. Salam, "Resonance Searches with an Updated Top Tagger," *JHEP* **06** (2015) 203, `arXiv:1503.05921 [hep-ph]`.

[362] **ATLAS** Collaboration, G. Aad *et al.*, "Measurements of top-quark pair differential cross-sections in the lepton+jets channel in *pp* collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector," `arXiv:1511.04716 [hep-ex]`.

[363] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of the differential cross-section of highly boosted top quarks as a function of their transverse momentum in $\sqrt{s} = 8$ TeV proton-proton collisions using the ATLAS detector," *Phys. Rev.* **D93** no. 3, (2016) 032009, `arXiv:1510.03818 [hep-ex]`.

[364] **CMS** Collaboration, V. Khachatryan *et al.*, "Measurement of the integrated and differential t-tbar production cross sections for high-pt top quarks in pp collisions at $\sqrt{s} = 8$ TeV," `arXiv:1605.00116 [hep-ex]`.

[365] M. Czakon, D. Heymes, and A. Mitov, "Dynamical scales for multi-TeV top-pair production at the LHC," `arXiv:1606.03350 [hep-ph]`.

[366] L. Berthier, M. Bjørn, and M. Trott, "Incorporating doubly resonant $W^{\pm}$ data in a global fit of SMEFT parameters to lift flat directions," `arXiv:1606.06693 [hep-ph]`.

[367] **ATLAS** Collaboration, G. Aad *et al.*, "Search for resonant top plus jet production in $t\bar{t}$ + jets events with the ATLAS detector in *pp* collisions at $\sqrt{s} = 7$ TeV," *Phys. Rev.* **D86** (2012) 091103, `arXiv:1209.6593 [hep-ex]`.

[368] **ATLAS** Collaboration, G. Aad *et al.*, "Search for resonances decaying into top-quark pairs using fully hadronic decays in *pp* collisions with ATLAS at $\sqrt{s} = 7$ TeV," *JHEP* **01** (2013) 116, `arXiv:1211.2202 [hep-ex]`.

[369] **CMS** Collaboration, S. Chatrchyan *et al.*, "Search for Z ' resonances decaying to $t\bar{t}$ in dilepton+jets final states in *pp* collisions at $\sqrt{s} = 7$ TeV," *Phys. Rev.* **D87** no. 7, (2013) 072002, `arXiv:1211.3338 [hep-ex]`.

[370] J. M. Butterworth, I. Ochoa, and T. Scanlon, "Boosted Higgs $\rightarrow b\bar{b}$ in vector-boson associated production at 14 TeV," *Eur. Phys. J.* **C75** no. 8, (2015) 366, `arXiv:1506.04973 [hep-ph]`.

[371] K. Joshi, A. D. Pilkington, and M. Spannowsky, "The dependency of boosted tagging algorithms on the event colour structure," *Phys. Rev.* **D86** (2012) 114016, `arXiv:1207.6066 [hep-ph]`.

[372] M. P. Rosello and M. Vos, "Constraints on four-fermion interactions from the $t\bar{t}$ charge asymmetry at hadron colliders," *Eur. Phys. J.* **C76** no. 4, (2016) 200, `arXiv:1512.07542 [hep-ex]`.

[373] F. Maltoni, D. Pagani, and I. Tsinikos, "Associated production of a top-quark pair with vector bosons at NLO in QCD: impact on t$\bar{t}$H searches at the LHC," *JHEP* **02** (2016) 113, arXiv:1507.05640 [hep-ph].

[374] O. Bessidskaia Bylund, F. Maltoni, I. Tsinikos, E. Vryonidou, and C. Zhang, "Probing top quark neutral couplings in the Standard Model Effective Field Theory at NLO in QCD," *JHEP* **05** (2016) 052, arXiv:1601.08193 [hep-ph].

[375] **ATLAS** Collaboration, M. Aaboud *et al.*, "Measurement of the t$\bar{t}$Z and t$\bar{t}$W production cross sections in multilepton final states using 3.2 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *Eur. Phys. J.* **C77** no. 1, (2017) 40, arXiv:1609.01599 [hep-ex].

[376] **CMS Collaboration** Collaboration, "Measurement of the top pair-production in association with a W or Z boson in pp collisions at 13 TeV," Tech. Rep. CMS-PAS-TOP-16-017, CERN, Geneva, 2016. https://cds.cern.ch/record/2205283.

[377] R. Contino, A. Falkowski, F. Goertz, C. Grojean, and F. Riva, "On the Validity of the Effective Field Theory Approach to SM Precision Tests," *JHEP* **07** (2016) 144, arXiv:1604.06444 [hep-ph].

[378] Y. Kiyo, A. Maier, P. Maierhofer, and P. Marquard, "Reconstruction of heavy quark current correlators at O(alpha(s)**3)," *Nucl. Phys.* **B823** (2009) 269–287, arXiv:0907.2120 [hep-ph].

[379] L. Chen, O. Dekkers, D. Heisler, W. Bernreuther, and Z.-G. Si, "Top-quark pair production at next-to-next-to-leading order QCD in electron positron collisions," *JHEP* **12** (2016) 098, arXiv:1610.07897 [hep-ph].

[380] M. Klute, R. Lafaye, T. Plehn, M. Rauch, and D. Zerwas, "Measuring Higgs Couplings at a Linear Collider," *Europhys. Lett.* **101** (2013) 51001, arXiv:1301.1322 [hep-ph].

[381] M. Vos, "Top physics beyond the LHC," arXiv:1701.06537 [hep-ex].

[382] M. Vos *et al.*, "Top physics at high-energy lepton colliders," arXiv:1604.08122 [hep-ex].

[383] M. S. Amjad *et al.*, "A precise characterisation of the top quark electro-weak vertices at the ILC," *Eur. Phys. J.* **C75** no. 10, (2015) 512, arXiv:1505.06020 [hep-ex].

[384] R. Röntsch and M. Schulze, "Probing top-Z dipole moments at the LHC and ILC," *JHEP* **08** (2015) 044, `arXiv:1501.05939 [hep-ph]`.

[385] B. Coleppa, M. Kumar, S. Kumar, and B. Mellado, "Measuring CP nature of top-Higgs couplings at the future Large Hadron electron collider," `arXiv:1702.03426 [hep-ph]`.

[386] Q.-H. Cao and B. Yan, "Determining V_{tb} at electron-positron colliders," *Phys. Rev.* **D92** no. 9, (2015) 094018, `arXiv:1507.06204 [hep-ph]`.

[387] S. T. Boogert and D. J. Miller, "Questions about the measurement of the e+ e- luminosity spectrum at a future linear collider," in *Linear colliders. Proceedings, International Workshop on physics and experiments with future electron-positron linear colliders, LCWS 2002, Seogwipo, Jeju Island, Korea, August 26-30, 2002*, pp. 509–516. 2002. `arXiv:hep-ex/0211021 [hep-ex]`. `http://alice.cern.ch/format/showfull?sysnb=2349913`.

[388] J. Fleischer, A. Leike, T. Riemann, and A. Werthenbach, "Electroweak one loop corrections for e+ e- annihilation into t anti-top including hard bremsstrahlung," *Eur. Phys. J.* **C31** (2003) 37–56, `arXiv:hep-ph/0302259 [hep-ph]`.

[389] J. Gao and H. X. Zhu, "Electroweak prodution of top-quark pairs in e+e- annihilation at NNLO in QCD: the vector contributions," *Phys. Rev.* **D90** no. 11, (2014) 114022, `arXiv:1408.5150 [hep-ph]`.

[390] W. Bernreuther, R. Bonciani, T. Gehrmann, R. Heinesch, T. Leineweber, P. Mastrolia, and E. Remiddi, "Two-Parton Contribution to the Heavy-Quark Forward-Backward Asymmetry in NNLO QCD," *Nucl. Phys.* **B750** (2006) 83–107, `arXiv:hep-ph/0604031 [hep-ph]`.

[391] J. Gao and H. X. Zhu, "Top Quark Forward-Backward Asymmetry in e^+e^- Annihilation at Next-to-Next-to-Leading Order in QCD," *Phys. Rev. Lett.* **113** no. 26, (2014) 262001, `arXiv:1410.3165 [hep-ph]`.

[392] K.-i. Hikasa, "Transverse Polarization Effects in e^+e^- Collisions: The Role of Chiral Symmetry," *Phys. Rev.* **D33** (1986) 3203.

[393] G. Moortgat-Pick *et al.*, "The Role of polarized positrons and electrons in revealing fundamental interactions at the linear collider," *Phys. Rept.* **460** (2008) 131–243, `arXiv:hep-ph/0507011 [hep-ph]`.

[394] M. S. Amjad, M. Boronat, T. Frisson, I. Garcia, R. Poschl, E. Ros, F. Richard, J. Rouene, P. R. Femenia, and M. Vos, "A precise determination of top quark

electro-weak couplings at the ILC operating at $\sqrt{s} = 500$ GeV," `arXiv:1307.8102` `[hep-ex]`.

[395] **CLICdp, CLIC** Collaboration, M. J. Boland *et al.*, "Updated baseline for a staged Compact Linear Collider," `arXiv:1608.07537 [physics.acc-ph]`.

[396] H. Abramowicz *et al.*, "Higgs Physics at the CLIC Electron-Positron Linear Collider," `arXiv:1608.07538 [hep-ex]`.

[397] J. Ellis, P. Roloff, V. Sanz, and T. You, "Dimension-6 Operator Analysis of the CLIC Sensitivity to New Physics," `arXiv:1701.04804 [hep-ph]`.

[398] M. E. Peskin and T. Takeuchi, "A New constraint on a strongly interacting Higgs sector," *Phys. Rev. Lett.* **65** (1990) 964–967.

[399] M. E. Peskin and T. Takeuchi, "Estimation of oblique electroweak corrections," *Phys. Rev.* **D46** (1992) 381–409.

[400] D. A. Ross and M. J. G. Veltman, "Neutral Currents in Neutrino Experiments," *Nucl. Phys.* **B95** (1975) 135–147.

[401] C. Grojean, E. E. Jenkins, A. V. Manohar, and M. Trott, "Renormalization Group Scaling of Higgs Operators and $\Gamma(h- > \gamma\gamma)$," *JHEP* **04** (2013) 016, `arXiv:1301.2588 [hep-ph]`.

[402] R. Barbieri, A. Pomarol, R. Rattazzi, and A. Strumia, "Electroweak symmetry breaking after LEP-1 and LEP-2," *Nucl. Phys.* **B703** (2004) 127–146, `arXiv:hep-ph/0405040 [hep-ph]`.

[403] **Gfitter Group** Collaboration, M. Baak, J. Cuth, J. Haller, A. Hoecker, R. Kogler, K. Mönig, M. Schott, and J. Stelzer, "The global electroweak fit at NNLO and prospects for the LHC and ILC," *Eur. Phys. J.* **C74** (2014) 3046, `arXiv:1407.3792 [hep-ph]`.

[404] **DELPHI** Collaboration, J. Abdallah *et al.*, "Measurement of the forward backward asymmetries of $e^+e^- \to Z \to b\bar{b}$ and $e^+e^- \to Z \to c\bar{c}$ using prompt leptons," *Eur. Phys. J.* **C34** (2004) 109–125, `arXiv:hep-ex/0403041 [hep-ex]`.

[405] M. H. Seymour, "Searches for new particles using cone and cluster jet algorithms: A Comparative study," *Z. Phys.* **C62** (1994) 127–138.

[406] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, "Jet substructure as a new Higgs search channel at the LHC," *Phys. Rev. Lett.* **100** (2008) 242001, `arXiv:0802.2470 [hep-ph]`.

[407] D. E. Kaplan, K. Rehermann, M. D. Schwartz, and B. Tweedie, "Top Tagging: A Method for Identifying Boosted Hadronically Decaying Top Quarks," *Phys. Rev. Lett.* **101** (2008) 142001, `arXiv:0806.0848 [hep-ph]`.

[408] J. Thaler and L.-T. Wang, "Strategies to Identify Boosted Tops," *JHEP* **07** (2008) 092, `arXiv:0806.0023 [hep-ph]`.

[409] L. G. Almeida, S. J. Lee, G. Perez, I. Sung, and J. Virzi, "Top Jets at the LHC," *Phys. Rev.* **D79** (2009) 074012, `arXiv:0810.0934 [hep-ph]`.

[410] Y. Cui, Z. Han, and M. D. Schwartz, "W-jet Tagging: Optimizing the Identification of Boosted Hadronically-Decaying W Bosons," *Phys. Rev.* **D83** (2011) 074023, `arXiv:1012.2077 [hep-ph]`.

[411] C. Anders, C. Bernaciak, G. Kasieczka, T. Plehn, and T. Schell, "Benchmarking an even better top tagger algorithm," *Phys. Rev.* **D89** no. 7, (2014) 074047, `arXiv:1312.1504 [hep-ph]`.

[412] D. Krohn, J. Thaler, and L.-T. Wang, "Jet Trimming," *JHEP* **02** (2010) 084, `arXiv:0912.1342 [hep-ph]`.

[413] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Techniques for improved heavy particle searches with jet substructure," *Phys. Rev.* **D80** (2009) 051501, `arXiv:0903.5081 [hep-ph]`.

[414] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches," *Phys. Rev.* **D81** (2010) 094023, `arXiv:0912.0033 [hep-ph]`.

[415] D. E. Soper and M. Spannowsky, "Finding physics signals with shower deconstruction," *Phys. Rev.* **D84** (2011) 074002, `arXiv:1102.3480 [hep-ph]`.

[416] J. Cogan, M. Kagan, E. Strauss, and A. Schwarztman, "Jet-Images: Computer Vision Inspired Techniques for Jet Tagging," *JHEP* **02** (2015) 118, `arXiv:1407.5675 [hep-ph]`.

[417] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, "Jet-images — deep learning edition," *JHEP* **07** (2016) 069, `arXiv:1511.05190 [hep-ph]`.

[418] P. T. Komiske, E. M. Metodiev, and M. D. Schwartz, "Deep learning in color: towards automated quark/gluon jet discrimination," `arXiv:1612.01551 [hep-ph]`.

[419] J. Barnard, E. N. Dawe, M. J. Dolan, and N. Rajcic, "Parton Shower Uncertainties in Jet Substructure Analyses with Deep Neural Networks," `arXiv:1609.00607 [hep-ph]`.

[420] L. G. Almeida, M. Backović, M. Cliche, S. J. Lee, and M. Perelstein, "Playing Tag with ANN: Boosted Top Identification with Pattern Recognition," *JHEP* **07** (2015) 086, `arXiv:1501.05968 [hep-ph]`.

[421] P. Baldi, K. Bauer, C. Eng, P. Sadowski, and D. Whiteson, "Jet Substructure Classification in High-Energy Physics with Deep Neural Networks," *Phys. Rev.* **D93** no. 9, (2016) 094034, `arXiv:1603.09349 [hep-ex]`.

[422] D. Guest, J. Collado, P. Baldi, S.-C. Hsu, G. Urban, and D. Whiteson, "Jet Flavor Classification in High-Energy Physics with Deep Neural Networks," *Phys. Rev.* **D94** no. 11, (2016) 112002, `arXiv:1607.08633 [hep-ex]`.

[423] P. Baldi, K. Cranmer, T. Faucett, P. Sadowski, and D. Whiteson, "Parameterized neural networks for high-energy physics," *Eur. Phys. J.* **C76** no. 5, (2016) 235, `arXiv:1601.07913 [hep-ex]`.

[424] G. Louppe, K. Cho, C. Becot, and K. Cranmer, "QCD-Aware Recursive Neural Networks for Jet Physics," `arXiv:1702.00748 [hep-ph]`.

[425] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions," *Nucl. Phys.* **B406** (1993) 187–224.

[426] S. D. Ellis and D. E. Soper, "Successive combination jet algorithm for hadron collisions," *Phys. Rev.* **D48** (1993) 3160–3166, `arXiv:hep-ph/9305266 [hep-ph]`.

[427] J. Thaler and K. Van Tilburg, "Identifying Boosted Objects with N-subjettiness," *JHEP* **03** (2011) 015, `arXiv:1011.2268 [hep-ph]`.

[428] J. Thaler and K. Van Tilburg, "Maximizing Boosted Top Identification by Minimizing N-subjettiness," *JHEP* **02** (2012) 093, `arXiv:1108.2701 [hep-ph]`.

[429] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft Drop," *JHEP* **05** (2014) 146, `arXiv:1402.2657 [hep-ph]`.

[430] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research* **12** (2011) 2825–2830.

[431] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, "An Introduction to PYTHIA 8.2," *Comput. Phys. Commun.* **191** (2015) 159–177, `arXiv:1410.3012 [hep-ph]`.

[432] **DELPHES 3** Collaboration, J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi, "DELPHES 3, A modular framework for fast simulation of a generic collider experiment," *JHEP* **02** (2014) 057, `arXiv:1307.6346 [hep-ex]`.

[433] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521** no. 7553, (5, 2015) 436–444.

[434] F. Chollet, "Keras." `https://github.com/fchollet/keras`, 2015.

[435] Theano Development Team, "Theano: A Python framework for fast computation of mathematical expressions," `arXiv:1605.02688 [cs.SC]`.

[436] W. Ding, "draw_convnet." `https://github.com/gwding/draw_convnet`, 2016.