

N° d'ordre NNT : 2017LYSE1142

THÈSE DE DOCTORAT DE L'UNIVERSITÉ DE LYON
ET DE L'UNIVERSITÉ CATHOLIQUE DE LOUVAIN

opérée en cotutelle au sein de
l'université Claude Bernard Lyon 1 et de
l'université catholique de Louvain

École Doctorale ED52
Physique et astrophysique de Lyon

Spécialité de doctorat : physique subatomique
Discipline : physique

Soutenue publiquement le 08/09/2017, par :
Nicolas Deutschmann

Precision calculations in effective theories for Higgs production

---

Devant le jury composé de :

Jean-Marc Gérard, Professeur, université catholique de Louvain	Président
Aude Gehrmann, Professeure, ETH Zürich	Rapporteuse
Gudrun Heinrich, Professeure, Max-Planck-Institut für Physik	Rapporteuse
Claude Duhr, Professeur, CERN	Examinateur
Giulia Zanderighi, Professeure, CERN	Examinatrice
Aldo Deandrea, Professeur, université de Lyon	Directeur de thèse
Fabio Maltoni, Professeur, université catholique de Louvain	Directeur de thèse

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université	**M. le Professeur Frédéric FLEURY**
Président du Conseil Académique	M. le Professeur Hamda BEN HADID
Vice-président du Conseil d'Administration	M. le Professeur Didier REVEL
Vice-président du Conseil Formation et Vie Universitaire	M. le Professeur Philippe CHEVALIER
Vice-président de la Commission Recherche	M. Fabrice VALLÉE
Directeur Général des Services	M. Alain HELLEU

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard	Directeur : M. le Professeur J. ETIENNE
Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux	Directeur : Mme la Professeure C. BURILLON
Faculté d'Odontologie	Directeur : M. le Professeur D. BOURGEOIS
Institut des Sciences Pharmaceutiques et Biologiques	Directeur : Mme la Professeure C. VINCIGUERRA
Institut des Sciences et Techniques de la Réadaptation	Directeur : M. le Professeur Y. MATILLON
Département de formation et Centre de Recherche en Biologie Humaine	Directeur : Mme la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies	Directeur : M. F. DE MARCHI
Département Biologie	Directeur : M. le Professeur F. THEVENARD
Département Chimie Biochimie	Directeur : Mme C. FELIX
Département GEP	Directeur : M. Hassan HAMMOURI
Département Informatique	Directeur : M. le Professeur S. AKKOUCHE
Département Mathématiques	Directeur : M. le Professeur G. TOMANOV
Département Mécanique	Directeur : M. le Professeur H. BEN HADID
Département Physique	Directeur : M. le Professeur J-C PLENET
UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur : M. Y.VANPOULLE
Observatoire des Sciences de l'Univers de Lyon	Directeur : M. B. GUIDERDONI
Polytech Lyon	Directeur : M. le Professeur E.PERRIN
Ecole Supérieure de Chimie Physique Electronique	Directeur : M. G. PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur : M. le Professeur C. VITON
Ecole Supérieure du Professorat et de l'Education	Directeur : M. le Professeur A. MOUGNIOTTE
Institut de Science Financière et d'Assurances	Directeur : M. N. LEBOISNE

Copyright information

Several figures in this thesis are subject to copyright:

Figure 2.2	© MissMJ (Wikimedia user)
Figure 2.3	© Particle Data Group
Figure 2.7	© CERN
Figure 2.6	© G. Watt
Figure 2.11	© LHC Higgs Cross Section Working Group
Figure 2.12	© CERN
Figure 2.13	© CERN
Figure 2.19	© Anastasiou, Duhr, Dulat, Herzog, Mistlberger
Figure 2.20	© CERN

Figure 2.19 was reproduced with kind permission from the authors. The others have been reproduced in conformity with their respective license requirements.

© Nicolas Deutschmann, 2017

This work is licensed under a Creative Commons "Attribution-NonCommercial-ShareAlike 4.0 International" license.



Precision calculations in effective theories for Higgs production

Nicolas Deutschmann

Acknowledgments

I would like to extend my warmest gratitude to the people who have made this thesis possible, and the four years I took preparing it enjoyable.

First and foremost I would like to thank my two PhD advisors, Aldo Deandrea and Fabio Maltoni, as well as my inofficial supervisor, Claude Duhr, for their continued support and guidance in my time as a PhD student in Louvain and Lyon. Despite not having worked as closely with Fabio as with Aldo and Claude during my PhD, I want to start by thanking him as it is he who introduced me to the world of particle physics by welcoming me to Louvain during my Master's for a summer internship six years ago; and he passed on to me his passion for the field. He has since then been a constant purveyor of cheerful support and opportunities to do exciting research. It is Fabio who proposed that I start working on radiative corrections and he rightfully perceived that I would find great pleasure in this field. I want to thank Aldo for teaching me so much about BSM phenomenology, model building and particle physics in general, for proposing interesting problems to study and for his encouragements and positivity, in particular through the tough realization that the initial research project that we developped would not bear fruit. Thanks also for leaving me complete freedom in my research interests and for supporting me through my change of topics from BSM physics to QCD. Finally, I want to thank Claude for introducing me to the world of QCD and precision calculations, his patience through the many hours he spent at a blackboard teaching me about loops and his support and understanding as I fumbled through making my first two-loop calculations. I cannot stress enough how much I have learned from and enjoyed our many discussions and how much I appreciate the effort he put in mentoring me.

I would also like to thank the jury that assessed my PhD defense and manuscript: Aude Gehrmann, Jean-Marc Gérard, Gudrun Heinrich, Giulia Zanderighi, together with Aldo, Claude and Fabio. The private defense was an unexpectedly enjoyable moment thanks to the very interesting discussions we had. They also provided insightful comments on this manuscript, which helped improve it significantly.

I wish to thank the people I have worked with during these four years, David Andriot, Giacomo Cacciapaglia, Alan Cornell, Federico Demartin, Thomas Flacke, Jong-Soo Kim and Eleni Vryonidou. Research is most often a team effort and I really enjoyed collaborating with all of them. Giacomo deserves a special thanks for the many interesting discussions we have had and for his committment to scientific rigor, which was instrumental in shaping the way I do research. I also want to extend a particular acknowledgement to Eleni for the many hours we spent bug-hunting when I started developping programs for my loop calculations. My codes work largely thanks to her.

I want to thank my coworkers, in Lyon, CP3 and CERN for the enjoyable environment they helped create during and outside working hours. I will not name names as the list would get very long, but I had a great time during lunches and coffee breaks. Thanks for the many debates, discussions and laughs.

I want to thank my family for their unwaivering faith in me, which has given me the confidence to push my passion for science to its accomplishement. Finally, and most importantly, I want to thank Mathilde for her support and love, which have been an immense source of confidence and happiness.

Publications

Part of this work has been accepted or is under review for publication

- N. Deutschmann, *Higgs production at NLO in the Standard Model Effective Field theory*, *in* Proceedings of the Kruger Workshop On Discovery Physics At The LHC 2016, Journal of Physics: Conference Series, Volume 878, conference 1.

- N. Deutschmann, C. Duhr, F. Maltoni, E. Vryonidou, *Gluon-fusion Higgs production in the Standard Model Effective Field Theory*, [arXiv:1708.00460] (submitted to JHEP).

During my thesis I have published other works of research not presented in this manuscript

- A. Deandrea, N. Deutschmann, *Multi-tops at the LHC*, JHEP 1408 (2014) 134, [arXiv:1405.6119]
  Short summary: we study the possibility that the production of BSM particles would give rise to exotic signatures where many top quarks are simultaneously produced. We work in the framework of simplified models and show that the measurements available after the LHC Run I already put constraints on these models.

- G. Cacciapaglia, A. Deandrea, N. Deutschmann, *Dark matter and localised fermions from spherical orbifolds?*, JHEP 1604 (2016) 083, [arXiv:1601.00081]
  Short summary: We study a class of models where the bosonic degrees of freedom of the SM are allowed to propagate in two extra-dimensions constructed from spherial orbifolds, while fermions are localized in 4D. We classified these models to find which can accommodate a Kaluza-Klein dark matter candidate. We explicitly studied the details of the minimal such construction and showed that it is excluded by the competition of resonance searches at the LHC and dark matter relic density constraints. We further argue that all other possible models are excluded as well.

- D. Andriot, G. Cacciapaglia, A. Deandrea, N. Deutschmann, D. Tsimpis, *Towards Kaluza-Klein Dark Matter on Nilmanifolds*, JHEP 1606 (2016) 169, [arXiv:1603.02289]
  Short summary: We consider a class of extra-dimensional models where the extra dimensions form a nilmanifold. We derived the scalar Kaluza-Klein spectrum on these spaces analytically and numerically and studied a toy model for scalar Kaluza-Klein dark matter with a localized SM.

- N. Deutschmann, T. Flacke, J. S. Kim, *Current LHC Constraints on Minimal Universal Extra Dimensions*, Phys. Lett. B771 (2017) 515-520, [arXiv:1702.00410]
  We update the constraints on the mUED model using the latest LHC searches for supersymmetric signatures implemented in `CheckMATE` and show that there is a nice complementarity of many searches that probes a large part of the parameter space of the model, improving significantly over previous limits.

Calculs de précisions dans des théories effectives pour la production du boson de Higgs

Résumé en français

La découverte du boson de Higgs en 2012 a marqué le début d'une nouvelle ère en physique des particules. Après cinq ans d'études, les caractéristiques de cette nouvelle particule confirment qu'elle se conforme aux prédictions du modèle standard de la physique des particules, bien que les incertitudes expérimentales permettent encore des déviations non-négligeables, qui pourraient être indicatrices de nouvelle physique. L'étude du boson de Higgs est donc actuellement tournée vers l'amélioration de la précision des mesures existantes et l'extension du corpus de données expérimentales à de nouveaux canaux de production et de désintégration afin de contraindre au mieux les possibles déviations.

Tester les propriétés du modèle standard en utilisant des mesures très précises requiert d'avoir à notre disposition des prédictions théoriques auxquelles les comparer. Ma thèse s'inscrit dans cette entreprise qui vise à améliorer la précision des calculs théoriques dans le modèle standard et au delà dans le but de fournir aux expérimentateurs les outils les plus performants possibles pour tester le modèle standard, contraindre les modèles de nouvelle physique, et extraire un maximum d'informations de découvertes éventuelles.

Après une introduction générale, ce manuscrit contient deux chapitres d'introduction, l'un décrivant le contexte physique et l'autre les techniques mathématiques qui ont permis le développement de cette thèse. L'introduction du contexte physique fait un survol des aspects les plus importants de la physique des particules au LHC en présentant d'abord le modèle standard, puis les spécificités des calculs de précisions en collisionneurs hadroniques, en particulier l'annulation des divergences infrarouges entre les corrections virtuelles et les émissions réelles. Nous abordons ensuite en plus de détails la physique du boson de Higgs, passant en revue son statut expérimental après le Run I du LHC et les méthodes d'ajustement qui ont servi aux mesures globales, avant de décrire la théorie effective du boson de Higgs, approximation du modèle standard dans la limite de la masse du quark top infinie, qui a permis de calculer la section efficace de production du boson de Higgs par fusion de gluons à une très grande précision. L'introduction du contexte physique se termine par l'introduction de la théorie effective du modèle standard, qui en est une extension permettant de décrire de manière universelle les effets indirects de nouvelle physique très lourde sur les interactions des particules élémentaires déjà connues.

La présentation des techniques mathématiques utilisées dans la thèse fait une revue des méthodes d'évaluation des diagrammes de Feynman à plusieurs boucles, qui sont les briques élémentaires des calculs perturbatifs aux ordres élevés. Nous présentons d'abord les techniques qui permettent de passer d'un grand nombre de diagrammes et d'intégrales à un petit nombre d'intégrales maîtresses, avant de décrire trois méthodes d'évaluation de celles-ci: la paramétrisation de Feynman qui permet l'évaluation directe des intégrales, l'expansion par régions et l'évaluation numérique par la méthode des secteurs. Le chapitre se termine par l'introduction des polylogarithmes multiples, fonctions spéciales essentielles à nos calculs, et à leur structure d'algèbre de Hopf, qui est un outil essentiel dans les calculs d'intégrales de Feynman.

Passés ces deux chapitres d'introduction, le manuscript présente les travaux originaux dévelopés au cours de cette thèse. Le chapitre 4 décrit l'extraction de la correction du couplage de Yukawa du quark

bottom dans la théorie effective du boson de Higgs par un calcul de correspondance à deux boucles entre cette théorie effective et le modèle standard. Cette correction représentait la pièce manquante pour l'amélioration de la prédiction de la section efficace de production du boson de Higgs en association avec deux quarks bottom, et nous pourrons l'utiliser à l'avenir pour réduire l'incertitude théorique de cette prédiction. Le calcul de correspondance est effectué par le biais d'un facteur de forme de la désintégration du boson de Higgs en deux quarks bottom. Les deux grandes parties du chapitre décrivent le calcul de ce facteur de forme. Nous effectuons d'abord un calcul exact dans la théorie effective du boson de Higgs en décrivant la décomposition de l'amplitude en termes d'intégrales maîtresses à une boucle, que nous évaluons par intégration directe après avoir introduit plusieurs techniques d'intégration dont une méthode itérative qui exploite la simplification des polylogarithmes par application du coproduit de manière itérée. Nous effectuons ensuite le calcul du même facteur de forme dans le modèle standard, où il est décrit par des diagrammes à deux boucles. Nous réduisons ceux-ci en termes d'intégrales maîtresses que nous calculons par intégration directe après expansion dans la limite de la masse du quark top infinie par la méthode des régions. Finalement, nous comparons les deux résultats pour fixer la valeur de la correction du couplage de Yukawa du quark bottom dans la théorie effective du boson de Higgs, qui prend une forme compacte.

Les deux chapitres finaux couvrent différents aspects du calcul de la correction au deuxième ordre de la section efficace de production d'un boson de Higgs par fusion de gluon dans la théorie effective du modèle standard, en présentant d'abord le calcul analytique des corrections virtuelles puis la phénoménologie du résultat. Le premier chapitre commence par la présentation des opérateurs de dimension six qui modifient la boucle du quark top dans la fusion de gluon et détaille les étapes qui ont permis d'évaluer analytiquement l'amplitude au premier ordre et sa correction virtuelle à deux boucles. Nous obtenons l'expression de l'amplitude en termes d'intégrales maîtresses identiques à celles utilisées dans le calcul de la section efficace de fusion de gluon dans le modèle standard, effectué en 2006, et utilisons ce résultat pour obtenir l'expression de la correction virtuelle de l'amplitude dans la théorie effective. Nous décrivons ensuite la structure des divergences de cette amplitude et appliquons sa renormalisation à une boucle. Nous utilisons ensuite la soustraction des divergences infrarouges pour isoler les divergences à deux boucles de l'amplitudes, dont la structure était inconnue avant notre calcul et fixons ainsi un des éléments de la matrice de renormalisation à deux boucles de la théorie effective du modèle standard. L'analyse des divergences infrarouges des contre-termes à une boucle permet d'expliquer l'apparition d'une divergence d'ordre trois dans l'amplitude non-renormalisée, fait inhabituel dans les calculs de corrections radiatives au second ordre. Ayant fixé entièrement la renormalisation de l'amplitude, nous pouvons exprimer sa contribution finie dans le formalisme de la soustraction de Catani-Seymour. Finalement, nous présentons la continuation analytique du résultat aux régions physique en décrivant le chemin suivi dans le plan complexe par le paramètre rationel en termes duquel sont exprimées les intégrales maîtresses, qui montre que des points de branchement apparaissent lorsque l'énergie passe le seuil de production du quark top et nous expliquons la méthode qui permet d'exprimer les polylogarithmes multiples dans le résutlat en termes de fonctions sans point de branchement et de logarithmes dont la continuation analytique est triviale.

Le second chapitre présente la combinaison du résultat à deux boucles avec le calcul automatique des corrections par émission réelles par le logiciel Madgraph5_aMC@NLO, qui permet l'intégration de la sec-

tion efficace évaluée au deuxième ordre et corrigée par l'inclusion de simulations de gerbes hadroniques. Le résultat de l'évaluation numérique est présenté et nous observons que la distribution différentielle en impulsion transverse peut discriminer l'effet de chaque opérateur dans la queue de la distribution. Nous discutons de l'impact des corrections radiatives sur le taux global de production du boson de Higgs ainsi que sur la section efficace différentielle en impulsion transverse. Ces corrections sont universelles pour les taux de production totaux, ce que l'on peut expliquer par la domination de la production au seuil, où les détails de l'interaction ne peuvent être résolus et chaque contribution prend la forme universelle d'une interaction de contact. Cette explication est confirmée par l'étude des corrections radiatives au niveau différentiel, où l'universalité est brisée à haute énergie. Nous finissons la discussion de la phénoménologie du processus par une rapide considération des effets éventuels d'opérateurs permettant à d'autres quarks d'apparaître dans la boucle de la fusion de gluons et concluons que seul le quark bottom pourrait contribuer de façon non-négligeable car d'autres mécanismes sont plus contraignants pour les quarks légers.

Contents

1 Introduction **3**

2 Physics at the Large Hadron Collider **7**
2.1 The standard model of particle physics 7
2.2 Making predictions for proton collisions 16
2.3 Higgs boson physics at the LHC 23
2.4 The standard model effective field theory 31

3 Multi-loop techniques **35**
3.1 Feynman diagrams, topologies and master integrals 35
3.2 Evaluating scalar integrals 42
3.3 Multiple polylogarithms 55

4 Matching the decay $H \to b\bar{b}$ between the SM and the HEFT **72**
4.1 Higgs-to-bottom decay in the Higgs Effective Field Theory 75
4.2 Higgs-to-bottom decay at two loops in the Standard Model 83
4.3 Extraction of the m_b/m_t correction to the bottom Yukawa 90

5 Gluon fusion in the SMEFT at two loops **91**
5.1 The gluon fusion process in the SMEFT 92
5.2 Calculation of the amplitude 94
5.3 Structure of UV and IR divergences 101
5.4 Renormalization group evolution of the parameters 104
5.5 Analytic continuation of the amplitude to the physical regions 105

6 Gluon fusion in the SMEFT at NLO **108**
6.1 Calculation setup 108
6.2 Theoretical uncertainties 109
6.3 Numerical results in the top sector 110
6.4 Discussion of the effects of light quarks 113

7 Conclusion **115**

A Shortcomings of Δ_Φ **117**

B Details of code implementation **119**
B.1 Generating diagrams with `QGRAF` and `qgraf-xml-parser` 119
B.2 Family reduction in practice: `SAIF` 120

C Two-loop master integrals in the $H \to b\bar{b}$ calculation **121**

D Details of the $gg \to H$ calculation **126**
D.1 Definition of the master integrals 126
D.2 Basis rotation at two loops 126
D.3 Renormalization scheme 130
D.4 Two-loop amplitudes 131

Bibliography **133**

1 Introduction

The Standard Model (SM) [1–3] is probably the most successful physical theory in history. It models the behavior of all the known elementary particles and their interactions through three of the four fundamental interactions and describes successfully most known physical phenomena that happen between human scales and the most microscopic scales ever probed, with the exception of gravitation. As a theory of the most elementary interactions between the elementary particles, it has endured impressive success in predicting the behavior of individual interactions between the building blocks of matter, which we have been probing for nearly a century in scattering experiments over six orders of magnitudes in energy, from the earliest Compton scattering experiments at $17\ \mathrm{keV}$ [4] to the record-breaking collisions at the LHC at $13\ \mathrm{TeV}$.

A major milestone in the recent history of the SM was the discovery of the Higgs boson of mass $125\ \mathrm{GeV}$ in 2012 [5, 6] by the ATLAS and CMS experiments at the LHC. The study of its properties is still ongoing, but we already have a rather comprehensive picture after the first run of the LHC [7] which shows that it behaves as predicted within the experimental uncertainties. The following years will be a critical period for the study of the SM as the experimental collaborations will push further the study of Higgs properties.

Despite the tremendous success of the SM in describing particle physics over decades of detailed scrutiny, we know that it cannot be the ultimate theory of nature because it fails to explain a number of experimental observations, the most obvious of which is the existence of gravitational interactions. Furthermore, there is now overwhelming evidence for the fact that neutrinos, which are massless in the SM, actually have a mass, which manifests itself in their flavor oscillations [8–10]. The existence of dark matter is also solidly grounded in observations at many scales, from galaxy rotation curves [11, 12], gravitational lensing in clusters [13, 14], cluster collisions [15] and the spectrum of the cosmic microwave background [16], which all seem to favor the existence of cold, weakly-interacting matter

which cannot be accounted for by the particle content of the SM.

As a result of these issues, much theoretical activity is devoted to developing models that go beyond the SM (BSM), predicting new particles and new structure to address the issues of the SM. Such models can be constrained or excluded by comparing their predictions to experimental measurements. In particular, many experimental analyses perform direct searches for new particles that could manifest themselves through a variety of signatures. Another avenue to constrain BSM physics is to consider precision observables: the effects of new physics could be subtle and manifest themselves through small deviations of processes involving SM particles. A striking example of this idea are the electroweak precision observables S, T, U [17] which were measured in Z boson production at LEP [18] and still provide a major source of constraints on BSM physics today. A number of precision measurements have been performed at the LHC on standard candle processes such as the production of jets [19, 20], electroweak boson [21, 22], or pairs of top quarks [23, 24]. Of course, testing the SM requires equally precise theoretical predictions, which have been the object of intense activity over the last years. Next-to-leading order (NLO) accuracy in Monte Carlo simulations have become the standard for generic applications since the publication of a series of automated codes [25–32], which built on major progress in our understanding of the analytic structure of one-loop amplitudes. Further orders in the perturbative expansion are much more difficult and require dedicated efforts for each result. As a result, NNLO calculations only exist for specific processes but have still been successfully performed for a number of key observables [33–39], while N^{3}LO calculations only exist for a handful of cases [40, 41]. While most NLO calculations have been automated, loop-induced processes also constitute a roadblock for automation due to the absence of universal reduction techniques such as those existing for one loop amplitudes.

While it is still relatively unconstrained compared to the iconic processes traditionnaly used for precision SM measurements, Higgs boson production has become a major aspect of SM physics at the LHC, and it already provides a number of observables that constrain possible new physics scenarios [42, 43]. The observation of many complementary production and decay modes has already permitted to constrain its interactions with other SM particles to a reasonable degree of precision and significant improvement is expected of the future runs of the LHC. On the theory side, the calculation of Higgs boson related observables is a very active subject, and we now have reliable predictions for the main production and decay processes in the SM [34, 35, 40, 44–50].

Making such predictions in the Higgs sector presents many challenges. In the SM, the main Higgs boson production mechanism is gluon fusion, a loop-induced process which includes a massive internal top quark line. As a result, it is very difficult to evaluate radiative corrections to this process and we only know it exactly at NLO accuracy [51, 52]. Thankfully, we know that we can make approximate predictions that reproduce the exact result very faithfully by considering the limit where the mass of the top quark goes to infinity. This approximation reduces the SM to an effective field theory (EFT) called the Higgs Effective Field Theory (HEFT) [53–55] in which the inclusion of more radiative corrections is possible, which has lead to the establishment of very precise predictions [40]. While originally developed with gluon fusion in mind, the HEFT provides a way to include the effects of top quark loops in other processes relevant for Higgs boson physics in a way that makes the evaluation of radiative corrections easier. The first project presented in this thesis was developed to provide a missing tool for such radiative corrections. We provided the first power-suppressed correction to the bottom-quark Yukawa

in the HEFT, which is relevant when considering the NLO correction to the associated production a pair of bottom quarks and a Higgs boson [44]. To this end, we performed the matching of a form factor for the $H \to b\bar{b}$ decay between the HEFT and the SM.

When considering BSM extensions of the SM Higgs sectors, another challenge is the large variety of models, which range from minimal extensions such as the two Higgs doublet model to very complex theories such as the Minimal Supersymmetric Standard Model [56, 57] or the many composite Higgs models [58, 59]. Using Higgs boson observables to constrain such models is a very fruitful approach, but systematically constraining the many BSM ideas available is a very involved task. As a result, the development of model independent ways of constraining new physics is a popular topic in modern phenomenology. The current approach used by experimental collaborations in the Higgs sector [7], proposed as a temporary framework by the Higgs cross section working group [45], is based on agnostic fits of SM predictions to experimental data. While it is properly universal, this approach provides constraints on quantities that are not well defined quantum-field-theoretical objects, making the connection to specific BSM scenarios complicated. Despite being less universal, an attractive alternative is the use of the SM effective field theory (SMEFT), which is the low energy limit of all BSM extensions whose new particles decouple at LHC energies [60–62]. This theory can describe all possible deviations to SM observables coming from heavy new physics, encoded as the Wilson coefficients of higher dimensional interactions between the SM fields. Because it is a framework defined in the language of quantum field theory, the relation between SMEFT parameters and physical observables can be established analytically and systematically improved by radiative corrections. The second project presented in this thesis is part of the ongoing effort to provide NLO predictions in the SMEFT in order to provide a large set of observables that could be used for a global fit of the SMEFT parameters. We considered the gluon fusion mechanism for Higgs boson production and calculated the NLO corrections to the total and differential cross sections. Because it is a loop-induced process, the NLO corrections involve evaluation of two-loop amplitudes, which made them impossible to evaluate them using automated NLO tools.

The outline of the thesis is as follows. In Chapter 2, we make a review of different aspects of LHC physics. We discuss the theoretical construction of the Standard Model and electroweak symmetry breaking, perturbative calculations for high-energy hadron scattering processes, experimental and theoretical features of Higgs boson physics, and finally describe the SMEFT. In Chapter 3, we introduce the mathematical tools for multi-loop calculations that have been relevant for our work. We first describe the techniques used to reduce the complexity of modern multi-loop amplitudes, focusing on the treatment of integrals based on integral families and the reduction to master integrals using integration-by-parts identities. We then make an overview of some Feynman integral evaluation techniques, describing the systematics of the Feynman parametrization and of the expansion by regions, which have been used extensively in this thesis, and providing a more schematic introduction to numerical integration methods, which were used as a cross check of our analytic result. We conclude this chapter by providing a review of multiple polylogarithms and their algebraic properties. After these introductory discussions, Chapter 4 presents our first original result: the evaluation of the power-suppressed correction to the bottom quark Yukawa in the SMEFT. We first motivate the need for this correction for the improvement of the prediction of Higgs boson prediction in association with a pair of bottom quarks and show how a form factor for the decay $H \to b\bar{b}$ can be used to extract this correction. We then

go over the calculation of this form factor in the SMEFT, which is a one-loop calculation. We derive the relevant one-loop integrals in a pedagogical way, introducing several techniques relevant for the direct integration of Feynman parametrized integrals and highlighting the usefulness of modern multiple polylogarithm techniques. We then describe the more complex calculation of the form factor in the SM, which requires the evaluation of two-loop Feynman diagrams. We compute the relevant two-loop master integrals and show explicitly how the expanded integrals in the SM are related to integrals appearing in the HEFT calculation. Finally, we compare the two results to extract the desired Wilson coefficient. The next two chapters, Chapter 5 and Chapter 6 describe our second original contribution: the evaluation of the NLO correction to the gluon fusion process in the SMEFT. The first chapter focuses on the evaluation of the virtual corrections to this process, which required the evaluation of two-loop Feynman diagrams. We first explain how the SMEFT modifies Higgs boson production by introducing the relevant dimension six operators that modify interactions between top quarks, gluons and Higgs bosons before describing the calculation of the amplitude at LO. We then describe the evaluation of bare amplitude at NLO and discuss the divergences of this amplitude. Using the known one-loop renormalization and infrared structures of the theory, we extract a previously unknown counterterm and use it to obtain the renormalized amplitude. We then discuss the running of the Wilson coefficients and the analytic continuation of the amplitude to the physical region, both of which will be used to study the phenomenology of the process. In Chapter 6, we describe the combination of the two loop virtual correction to the amplitude with the real emission diagrams generated by `MadGraph5_aMC@NLO` to produce physical predictions for gluon fusion in the HEFT. We then address the issue of the theoretical uncertainties of our predictions related to the missing orders in the perturbative QCD expansion and the power suppressed terms in the SMEFT expansion. We present results for the total cross section and kinematic distributions as a function of the Wilson coefficients and discuss the effects of the NLO corrections on these observables. Finally, we consider the possibility to extend our prediction to include the effects of light quarks in the loop and discuss their phenomenological relevance. We provide concluding remarks in Chapter 7.

2

Physics at the Large Hadron Collider

2.1 The standard model of particle physics

2.1.1 Theoretical structure

The SM is a quantum field theory (QFT), defined by its set of quantum fields, their symmetry properties, and an action, S_{SM}, which encodes their behavior. The action can be entirely constructed from the symmetry properties of the fields and the constraint of renormalizability and this section will outline the principles of this construction and their physical consequences. The action is invariant under the following symmetry group:

$$G_{SM} = \mathbf{P}_4 \otimes U(1)_Y \otimes SU(2)_L \otimes SU(3)_c, \tag{2.1}$$

where $\mathbf{P}_4$ is the 4-dimensional Poincaré group[1] which describes the invariance of the theory under special relativity transformations and the other groups are gauge groups, describing internal symmetries of the fields. The labels of the gauge groups are there to indicate the role each group plays in the SM: $U(1)_Y$ is the hypercharge group, whose charge is $\hat{y}$, $SU(2)_L$ couples only to left-handed fermions and the Higgs boson, and $SU(3)_c$ is the color group.

[1] Note that in practice we will work in the context of dimensional regularization, where we will compute all quantities in $d = 4 - 2\epsilon$ dimensions and analytically continue observables in the limit $\epsilon \to 0$.

The Standard Model classical action S_{SM} can be compactly expressed as

$$\begin{aligned} S_{\text{SM}} = \int_{-\infty}^{+\infty} d^d x \Big(& -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} - \frac{1}{4} W_I^{\mu\nu} W^I_{\mu\nu} - \frac{1}{4} G_a^{\mu\nu} F^a_{\mu\nu} \\ & + i\bar{L} \not{D} L + i\bar{e} \not{D} e + i\bar{Q}_L \not{D} Q_L + i\bar{\tilde{u}}_R \not{D} \tilde{u}_R + i\bar{d}_R \not{D} d_R \\ & + \left(\bar{L}\phi^c\right) Y_l e + \bar{e} Y_l^\dagger \left(\phi^{c\dagger} L\right) \\ & - \left(\bar{Q}_L \phi\right) Y_d d_R - \bar{d}_R Y_d^\dagger \left(\phi^\dagger Q_L\right) \\ & - \left(\bar{Q}_L \phi^c\right) Y_u \tilde{u}_R - \bar{\tilde{u}}_R Y_u^\dagger \left(\phi^{c\dagger} Q_L\right) \\ & + |D\phi|^2 - \mu^2 \phi^\dagger \phi + \lambda \left(\phi^\dagger \phi\right)^2 \Big) \end{aligned} \tag{2.2}$$

Where $F, W, G, L, e, Q_L, \tilde{u}_R, d_R, \phi$ are functions of spacetime that take their value in some irreducible representation of $P_4 \otimes U(1)_Y \otimes SU(2)_L \otimes SU(3)_c$.

Poincaré invariance is automatically satisfied by the fact that the action is written as the integral of a Lagrangian

$$S_{\text{SM}} = \int_{-\infty}^{+\infty} d^4 x \mathcal{L}_{\text{SM}} \tag{2.3}$$

where each term of the Lagrangian is *local* and is a Lorentz scalar. The constraint of renormalizability means that the set of operators in $\mathcal{L}_{\text{SM}}$ is stable under quantum corrections, which is equivalent to the absence of terms of the form $g\mathcal{O}$, where g is a parameter and $\mathcal{O}$ contains only fields and derivatives and is of dimension > 4.

While Poincaré symmetry is manifest, the invariance of $\mathcal{L}_{\text{SM}}$ under the gauge group is non-trivially realized. Gauge symmetry requires the invariance under *local* transformations of the gauge group. For example, spin-$1/2$ fields transform linearly, which means that these fields are vectors of some irreducible representation of this gauge group and that for any function $U(x)$ from spacetime to this irreducible representation, the transformation

$$\psi(x) \to U(x)\psi(x) \tag{2.4}$$

leaves the Lagrangian invariant. Using the generators K^a of the representation, this transformation can be re-expressed as

$$\psi(x) \to \exp\left(i\alpha_a(x) K^a\right) \psi(x), \tag{2.5}$$

for some function $\alpha_a(x)$. This, in particular, means that the kinetic term cannot enter the Lagrangian alone: under such a transformation

$$i\bar{\psi}(x) \not{\partial} \psi(x) \to i\bar{\psi}(x) \not{\partial} \psi(x) + i\bar{\psi}(x) \left(U^\dagger \not{\partial} U\right) \psi(x). \tag{2.6}$$

This is a consequence of the fact that the field ψ takes its values in the tangent bundle of a fiber bundle whose base is Minkowski space and whose fiber is the gauge group: the derivative needs to contain a connection term that describes the geometry of the fiber. In that language, the invariance under gauge transformations corresponds to the invariance under reparametrization of the fiber. The covariant

derivative should therefore be

$$D_\mu = \partial_\mu - iC_\mu(x), \tag{2.7}$$

where C_μ is the connection and takes its values in the tangent space of the gauge group, *i.e.*, the Lie algebra of the gauge group, and transforms in the following way:

$$C_\mu(x) \to U^\dagger(x)C_\mu(x)U(x) + U^\dagger(x)\partial_\mu U(x) \tag{2.8}$$

which leaves $\bar{\psi}\not{D}\psi$ invariant under gauge transformations. It is often convenient to decompose $C_\mu(x)$ in the basis provided by the generators of the Lie algebra:

$$C_\mu(x) = C_{\mu a}(x)G^a. \tag{2.9}$$

The exact form of the covariant derivative is implicitly dependent on the representation of the field it acts on, and provides a generic way to construct gauge-invariant kinetic terms for fermions. The kinetic terms for the fermion fields $L, e, Q_L, \tilde{u}_R, d_R$ are built exactly in this way in the second line of Equation 2.2. These terms saturate the complete list of symmetric renormalizable operators involving only fermion fields and the coupling between the fermions and the gauge connection $C_\mu(x)$.

The names and representations of the SM fermions are described in Table 2.1. Each fermion is actually a three-component field, with each component called a *flavor* or a *family*, which will be distinguished by the structure of the Yukawa matrices Y_i, which are 3×3 matrices in flavor space.

Name	Symbol	$U(1)_Y \otimes SU(2)_L \otimes SU(3)_c$
lepton doublet	$L_e = \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}$	$-\mathbf{1} \otimes \mathbf{2} \otimes \mathbf{1}$
charged lepton singlet	e_R	$-\mathbf{2} \otimes \mathbf{1} \otimes \mathbf{1}$
quark doublet	$Q_L = \begin{pmatrix} \tilde{u}_L \\ d_L \end{pmatrix}$	$\mathbf{1/3} \otimes \mathbf{2} \otimes \mathbf{3}$
up quark singlet	$\tilde{u}_R$	$\mathbf{4/3} \otimes \mathbf{1} \otimes \mathbf{3}$
down quark singlet	d_R	$-\mathbf{2/3} \otimes \mathbf{1} \otimes \mathbf{3}$

Table 2.1: The fermion fields of the SM. Each field has three components which represent the three flavors of each type of SM fermion.

In gauge theories such as the SM, the geometry of the gauge fiber bundle is a dynamic object. As a result, the gauge connections $C_{\mu a}(x)$ are promoted to dynamic fields, which will correspond to the mediators of the fundamental interactions. Because our gauge group is a semi-simple Lie group, it is useful to use that structure to decompose its Lie algebra as a direct sum $\mathfrak{u}(1) \oplus \mathfrak{su}(2) \oplus \mathfrak{su}(3)$ and to

choose an adapted basis:

$$\begin{aligned} &\hat{y} \oplus \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ &0 \oplus \tau^I \oplus \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ &0 \oplus \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \oplus T^a, \end{aligned} \tag{2.10}$$

where the τ^a and T^a are respectively the generators of the relevant representation of $SU(2)$ and $SU(3)$. We will systematically drop all trivial parts of the algebra and use $\hat{y}$, τ and T to mean the matrices with $0_{n\times n}$ blocs. To reflect the fact that the gauge group describes several interactions, we introduce the coupling constants g, g' and g_s corresponding to each interaction and define the following fields

$$C_\mu(x) = g'\frac{\hat{y}}{2}B_\mu(x) + gW_{\mu I}\tau^I + g_s G_{\mu a}T^a. \tag{2.11}$$

The B and W_I fields are called the electroweak gauge fields and G the gluon field. The gluons, are particles whose effect is seen in colliders and were discovered in 1979, while the electroweak bosons have generic names because they will be recombined under electroweak symmetry breaking to form the other known vector bosons, the photon γ, definitely established in 1923, and the W and Z bosons, discovered in 1983.

The basic building blocks for the gauge sector of the SM Lagrangian are the field strengths, which are the simplest objects that transform linearly under the gauge group:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\mu A_\nu, \tag{2.12}$$

$$W^I_{\mu\nu} = \partial_\mu W^I_\nu - \partial_\mu W^I_\nu + g\epsilon^{IJK}W_{J\mu}W_{K\nu}, \tag{2.13}$$

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\mu G^a_\nu + g_s f^{abc}G_{b\mu}G_{c\nu}. \tag{2.14}$$

The kinetic terms appearing in the first line of Equation 2.2 are constructed of the contraction of each field strength with itself. For the non-abelian fields W and G, another structure is however possible: $\epsilon^{\mu\nu\rho\sigma}W^I_{\mu\nu}W_{I\rho\sigma}$ and $\epsilon^{\mu\nu\rho\sigma}G^a_{\mu\nu}G_{a\rho\sigma}$, while the corresponding term for A yields a trivial contribution to the action because it can be shown to be a total derivative. Similarly, the term $\epsilon^{\mu\nu\rho\sigma}W^I_{\mu\nu}W_{I\rho\sigma}$ can be safely eliminated from the Lagrangian by some transformations of the fields [63] that leaves the classical Lagrangian unchanged. On the other hand, the most general gauge symmetric Lagrangian should contain a term of the form $\theta\epsilon^{\mu\nu\rho\sigma}G^a_{\mu\nu}G_{a\rho\sigma}$. It can be shown that such a term would induce an electric dipole moment in neutrons, which has never been observed experimentally and constrains $\theta < 10^{-11}$ [64], which explains why it is not included in the SM Lagrangian. It is not understood why this term is zero or extremely supressed in the SM, and this issue is referred to as the strong-CP problem, which can only be solved by extensions of the Standard Model such as the Peccei-Quinn model [65]. Except for this operator, the first line of the SM Lagrangian contains all the renormalizable operators involving only gauge bosons preserving its gauge symmetries.

The final component of the Standard Model is the Higgs field ϕ, which was introduced by Brout, Englert and Higgs [66, 67] and is the source of the breaking of the $U(1) \otimes SU(2)$ part of the gauge symmetry. This field is a Lorentz scalar in the $1 \otimes 2 \otimes 1$ representation of the gauge group. There are three possible renormalizable gauge invariant operators that involve the Higgs boson, which are all included in the last line of Equation 2.2. The last two terms $-\mu^2 \left(\phi^\dagger \phi\right) + \lambda \left(\phi^\dagger \phi\right)^2$ are referred to as the Higgs potential and we will discuss in the next section how it provides a mechanism to generate masses for the gauge bosons and the fermions.

Finally, the Yukawa sector is composed of the 4th and 5th lines of Equation 2.2 and is composed of the last 4 possible gauge invariant terms. In each of these terms, the Higgs field is contracted with a left-handed fermionic $SU(2)$ doublet and a right handed field. A priori, the Yukawa matrices appearing in the Lagrangian are generic matrices, but we can use the symmetries of the Lagrangian to constrain them. Indeed, we can always rotate our fermion fields in flavor space

$$\begin{array}{ll} L \to U_L L & e_R \to U_e e_r \\ Q_L \to V_q Q_L & \tilde{u}_R \to V_u \tilde{u}_R \quad d_R \to V_d d_R, \end{array} \tag{2.15}$$

where the V_i and U_i are orthogonal matrices acting on flavor. In the lepton sector, we can use these transformations to diagonalize the leptonic Yukawa matrix Y_l, which also leaves the kinetic term unchanged. As a result, we can consider that

$$Y_l = \begin{pmatrix} y_e & 0 & 0 \\ 0 & y_\mu & 0 \\ 0 & 0 & y_\tau, \end{pmatrix} \tag{2.16}$$

where the indices in the eigenvalues indicate that the flavors correspond respectively to the electron, the muon and the tau and the eigenvalues are real. On the other hand in the quark sector, Y_u and Y_d cannot be simultaneously diagonalized in general. The convention in the Standard Model is to diagonalize Y_d and leave the mixing in Y_u. We can choose V_q, V_d and V_u such that

$$Y_d = V_q \hat{Y}_d V_d^T, \qquad Y_u = V \hat{Y}_u V_u^T, \tag{2.17}$$

where $\hat{Y}_d$ and $\hat{Y}_u$ are diagonal matrices and V is an orthogonal matrix. Let us define $\mathcal{V} = V_q^T V$, we can always write the quark Yukawa terms in the Lagrangian as

$$-\left(\bar{Q}_L \phi\right) \hat{Y}_d d_R - \left(\bar{Q}_L \phi^c\right) \mathcal{V} \hat{Y}_u \tilde{u}_R + \text{h.c.} \tag{2.18}$$

The leftover matrix $\mathcal{V}$ that we cannot rotate away is the CKM matrix, which has the physical consequence of allowing flavor-changing charged currents in the Standard Model.

2.1.2 Electroweak symmetry breaking

The structure of the Standard Model in its manifestly gauge-symmetric form in Equation 2.2 seems to have many different features from real-life physics. Indeed, no fermion has a mass term in the Lagrangian because the gauge structure forbids direct linear couplings between left- and right-handed

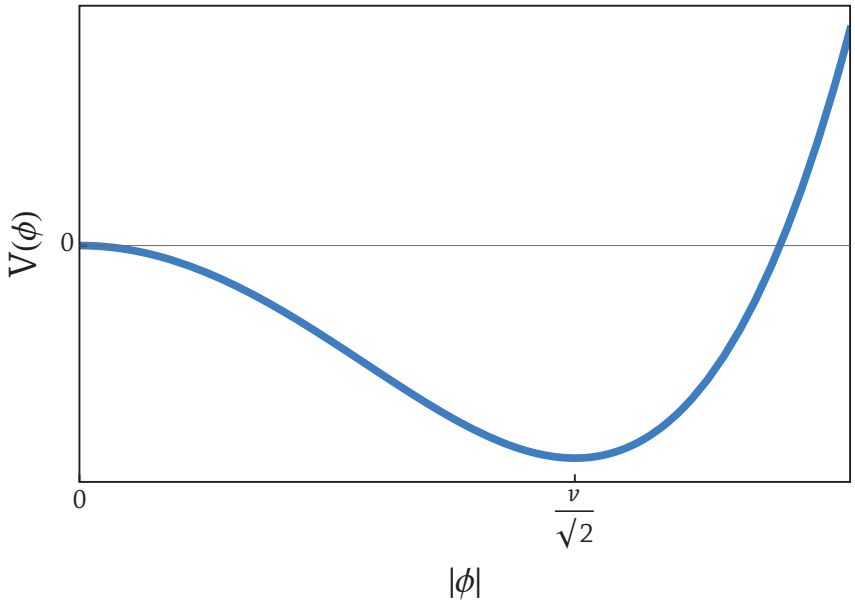


Figure 2.1: The Higgs potential as a function of $|\phi| = \sqrt{\phi^\dagger \phi}$. The minimum of the potential is at $|\phi| = v$.

fields. Futhermore, while we have observed the existence of massive vector bosons, gauge symmetry also forbids the existence of mass terms. This apparent discrepancy is due to the fact that the SM undergoes electroweak symmetry breaking, which is a process in which the vacuum breaks the symmetry, allowing for mass terms to appear. This is an essential feature of the SM as theories with massive gauge bosons are generally non-renormalizable unless their mass is generated through spontaneous symmetry breaking [68].

Spontaneous symmetry breaking is realized because of the structure of the Higgs potential:

$$V(\phi) = -\mu^2 \left(\phi^\dagger \phi\right) + \lambda \left(\phi^\dagger \phi\right)^2, \tag{2.19}$$

which can be represented graphically as a function of $|\phi| = \sqrt{\phi^\dagger \phi}$ as in Figure 2.1.

This representation makes it manifest that this potential has a minimum away from the usual $\phi = 0$. As a result, the physical ground state should correspond to a situation where $\langle |\phi| \rangle = \frac{v}{\sqrt{2}}$ where the minimum of the potential is reached. By minimizing Equation 2.19, we find that

$$v^2 = \frac{\mu^2}{\lambda}, \tag{2.20}$$

and our perturbative expansion should be taken around such a ground state. Because ϕ is a $SU(2)$ doublet, we can parametrize it in the following way

$$\phi(x) = \exp(i\pi(x) \cdot \tau) \begin{pmatrix} 0 \\ \frac{v + h(x)}{\sqrt{2}} \end{pmatrix}, \tag{2.21}$$

where the τ are the generators of $SU(2)$ and the dynamical degrees of freedom are the Goldstone fields

$\pi_i(x)$ and the physical higgs field $h(x)$. This choice corresponds to a vacuum in which

$$\langle\phi\rangle = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}. \tag{2.22}$$

There are infinitely many degenerate vacua as any $SU(2)$ transformation of this state provides a different state which minimizes the potential: indeed, the theory itself is still $SU(2)$ invariant.

Let us outline the main consquences of the spontaneous symmetry breaking on the SM Lagrangian. To this end, we will work in the *unitary gauge*, in which we perform a local $SU(2)$ transformation to eliminate the dependence on the fields $\pi(x)$:

$$\phi(x) \to \exp(-i\pi(x)\cdot\tau)\phi(x) = \begin{pmatrix} 0 \\ \frac{v+h(x)}{\sqrt{2}} \end{pmatrix}, \tag{2.23}$$

which also transforms all the other states. The general form of the Lagrangian in Equation 2.2 is left unchanged, but this simplifies the expansion of the Higgs fields into its components greatly. In this gauge, the potential is reduced to

$$V(h) = -\lambda\frac{v^2}{4} + v^2\lambda h^2 + v\lambda h^3 + \frac{\lambda}{4}h^4. \tag{2.24}$$

As a result, it appears that the field h describes a neutral scalar particle with mass $m_H = v\sqrt{2\lambda}$, the Higgs boson. Let us now work out what happens to the covariant derivative of the Higgs field by expanding explicitly the generators in the $SU(2)$ field W:

$$|D\phi|^2 = \frac{1}{2}\left(\partial_\mu h + (h+v)\left(-igB_\mu - ig'W_\mu\right)\right)\left(\partial^\mu h + (h+v)\left(+igB^\mu + ig'W^{\mu\dagger}\right)\right) \tag{2.25}$$

$$\begin{aligned} &= \frac{1}{2}\partial_\mu h\partial^\mu h + \frac{1}{8}\left(g'B_\mu - gW^3_\mu\right)\left(g'B^\mu - gW^{3\mu}\right)(v+h)^2 \\ &\quad + \frac{g^2}{8}\left(W^1_\mu + iW^2_\mu\right)\left(W^{1\mu} - iW^{2\mu}\right)(v+h)^2 \end{aligned} \tag{2.26}$$

$$= \frac{1}{2}\partial_\mu h\partial^\mu h + \left(g^2+g'^2\right)\frac{v^2}{8}Z_\mu Z^\mu + \frac{g^2v^2}{4}W^+_\mu W^\mu_- + \text{interactions}. \tag{2.27}$$

As a result, the Higgs field covariant derivative generates the kinetic term for h and masses for two combinations of gauge bosons. The Z field is a combination of the B and the W_3 fields to form a neutral massive vector field and W_1 and W_2 combine to form a charged massive vector field, $W^\pm$. Out of the four electroweak gauge boson, there is therefore a last combination which remains massless, which corresponds to the photon field A:

$$Z_\mu = \frac{g'B_\mu - gW^3_\mu}{\sqrt{g^2+g'^2}},\ A_\mu = \frac{gB_\mu + g'W^3_\mu}{\sqrt{g^2+g'^2}},\ W^\pm_\mu = \frac{W^1_\mu \mp iW^2_\mu}{\sqrt{2}}. \tag{2.28}$$

The photon field is associated to a leftover $U(1)$ symmetry, leaving a conserved quantum number, the

charge $\mathcal{Q}$. The charge operator depends on the hypercharge and the third generator of $SU(2)$[2]:

$$\mathcal{Q} = \frac{y}{2} + \tau^3. \tag{2.29}$$

Finally, let us outline the effect of symmetry breaking on the fermion sector. The only affected part of the Lagrangian is the Yukawa sector. Remember that we had the Yukawa part of the Lagrangian

$$\mathcal{L}_Y = -\left(\bar{L}\phi\right)\hat{Y}_l e_R - \left(\bar{Q}_L\phi\right)\hat{Y}_d d_R - \left(\bar{Q}_L\phi^c\right)\mathcal{V}\hat{Y}_u\tilde{u}_R + \text{h.c.} \tag{2.30}$$

$$= -\sum_{f=e,\mu,\tau}\frac{y_f v}{\sqrt{2}}\bar{e}_L^f e_R^f - \sum_{f=d,s,b}\frac{y_f v}{\sqrt{2}}\bar{d}_L^f d_R^f - \bar{\tilde{u}}_L\mathcal{V}\frac{v\hat{Y}_u}{\sqrt{2}}\tilde{u}_R + \text{h.c} \tag{2.31}$$

Let us define the physical up-quark states:

$$\begin{pmatrix} u_L \\ c_L \\ t_L \end{pmatrix} = \mathcal{V}^\dagger\tilde{u}_L, \tag{2.32}$$

$$\begin{pmatrix} u_R \\ c_R \\ t_R \end{pmatrix} = \tilde{u}_R. \tag{2.33}$$

Using this notation, the Yukawa sector is entirely diagonalized

$$\mathcal{L}_Y = -\sum_{f=e,\mu,\tau}\frac{y_f v}{\sqrt{2}}\bar{e}_L^f e_R^f - \sum_{f=d,s,b}\frac{y_f v}{\sqrt{2}}\bar{d}_L^f d_R^f - \sum_{f=u,c,t}\frac{y_f v}{\sqrt{2}}\bar{u}_L^f u_R^f + \text{h.c.} \tag{2.34}$$

This change leaves the kinetic term for the up quarks diagonal, so this choice of basis provides the correct definition for the physical fields in the SM.

After electroweak symmetry breaking, the physical spectrum of fundamental particles can be summarized as in Figure 2.2.

Every one of these particles has been observed experimentally in a span of 115 years, from the discovery of the electron in 1887 to that of the Higgs boson at the LHC in 2012 [5, 6].

2.1.3 Quantum chromodynamics

The theory of the strong interaction, mediated by gluons, is called quantum chromodynamics (QCD). The QCD Lagrangian, which is contained in the SM Lagrangian after electroweak symmetry breaking, takes a much simpler form than the electroweak sector:

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}G^a_{\mu\nu}G_a^{\mu\nu} + \bar{q}^f\left(i\not{D} - m_f\right)q^f. \tag{2.35}$$

This Lagrangian is however deceiptively simple, as the phenomenology of QCD is very rich and changes dramatically across energy scales. At high energy, taking a perturbative expansion can be used to predict physical phenomena reliably. As a result, it makes sense to describe hard physical processes in terms

[2]The generator is representation-dependent

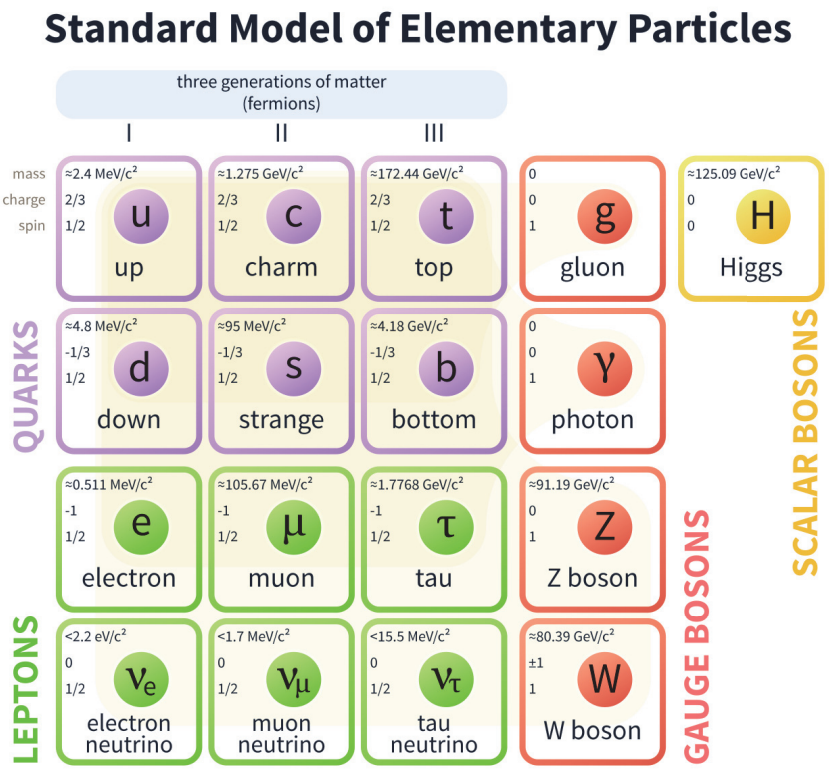


Figure 2.2: The physical spectrum of the SM.

of interactions between quarks and gluons and to express physical observables in powers of the $SU(3)$ coupling constant g_s.

Low-energy QCD is however very different due to the phenomenon of *confinement*. Indeed, at energies lower than the confinement scale $\Lambda_{\text{QCD}} \approx 200\ \text{MeV}$, quarks and gluons are confined into colorless states called hadrons. Hadrons come in a wide variety and hundreds of different states have been observed experimentally since the first discovery of a hadron, the proton, in 1917. Describing the behavior of hadrons from first principles is extremely complicated because of the loss of perturbativity for energies below Λ_{QCD}. Indeed, as shown in Figure 2.3, the strong coupling constant $\alpha_s = g_s^2/(4\pi)$ has the peculiar feature of growing as the energies goes from infinity to Λ_{QCD}, around which the theory becomes strongly coupled and we cannot use perturbative expansions to make predictions. A number of effective theories exist that describe low energy QCD in a bottom-up approach, but predicting hadron behavior from first principles has for long been a very tough nut to crack. The current best approach is lattice QCD, which attempts to evaluate the path integral by discretizing spacetime and has become a powerful tool.

A crucial point for our ability to use perturbative QCD is that, at high energies, the quarks and gluons inside hadrons, collectively called partons, have individual interactions that *factorize* from the nonperturbative dynamics of the hadron. Indeed, the high energy interactions of hadrons can be predicted as a decoherent sum over the interactions of their components, meaning that we can separate the hard interactions of elementary particles from the hadronic physics which happens at a different scale. In the next section, we will discuss how this phenomenon is used to make perturbative QCD predictions

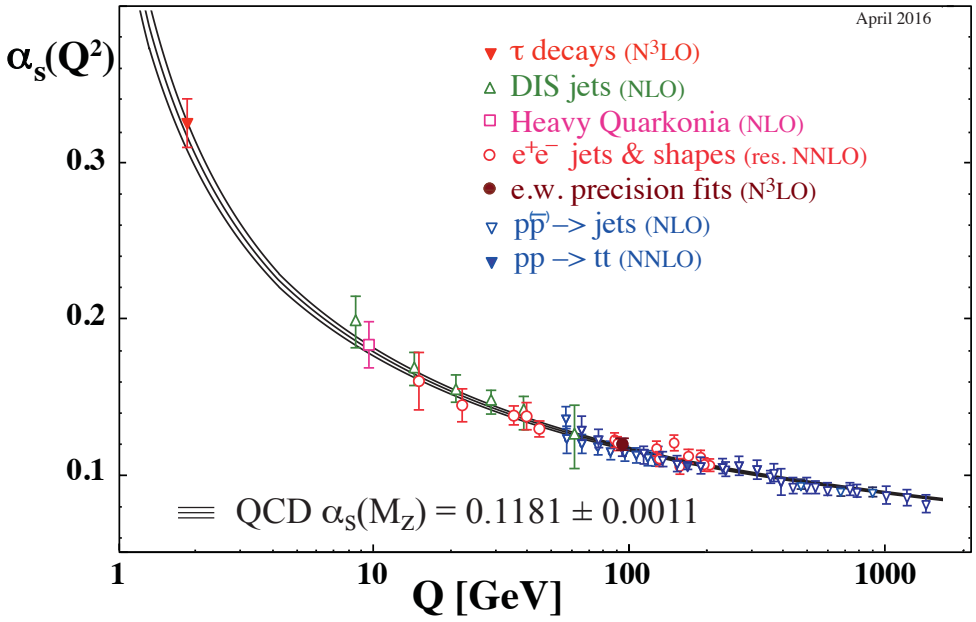


Figure 2.3: Evolution of the strong coupling α_s with the energy scale Q [69]. The theory becomes nonperturbative below $1\ \mathrm{GeV}$

at hadron colliders.

2.2 Making predictions for proton collisions

2.2.1 Brief overview of perturbative QFT

Most modern particle physics experiment are scattering experiments, in which one studies the collisions of subatomic particles to determine how they interact. Predicting the outcome of such experiment using QFT is done by calculating *transition amplitudes*, whose norm is the probability for a given physical process to occur. The quantum operator describing general transitions between asymptotic states (states where particles are well separated and do not interact anymore) is called the S-matrix, providing a compact definition of a transition amplitude $\mathcal{A}_{i\to f}$ from an initial state i to a final state f:

$$\mathcal{A}_{i\to f} = \langle f \,|\, S \,|\, i \rangle \,. \tag{2.36}$$

Determining the S-matrix in general for a QFT is a very difficult task and we are most often content with finding approximate solutions using perturbation theory, in which we expand around a free theory in powers of the interactions couplings. For example, in the case of QCD, the strength of all interactions is controlled by the coupling g_s and high-energy observables such as scattering cross sections can be expressed as an asymptotic series in the strong coupling

$$\alpha_s = \frac{g_s^2}{4\pi}. \tag{2.37}$$

In general, finding individual elements of the S-matrix as an expansion in powers of a coupling is by itself a difficult task. The canonical approach is to use the fact that, by the LSZ theorem [70], elements of the S-matrix are related to time-ordered correlation functions. By the Wick theorem [71], such correlation functions can be expressed as linear combinations of products of two-point functions, and the expression of terms in the perturbative expansion is reduced to a combinatorial exercise. Modern-day calculations usually bypass these steps by using a pictorial approach based on Feynman diagrams. Feynman diagrams are constructed using Feynman rules, which can be determined directly from the Lagrangian of a theory in an algorithmic way and nowadays, several tools exist that can perform this task automatically [72]. As an example, the Feynman rules of Quantum Chromodynamics are described in Figure 2.4.

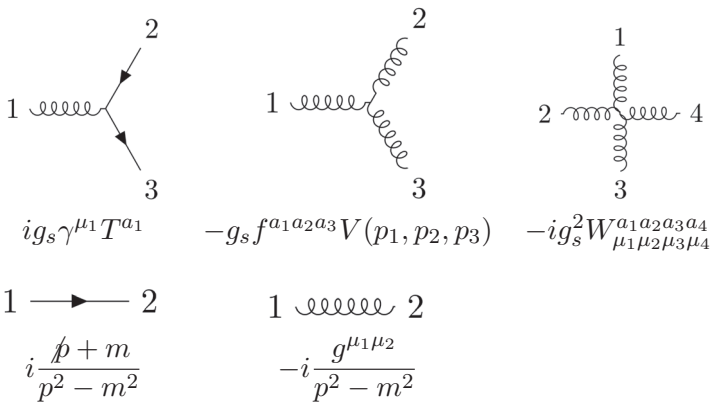


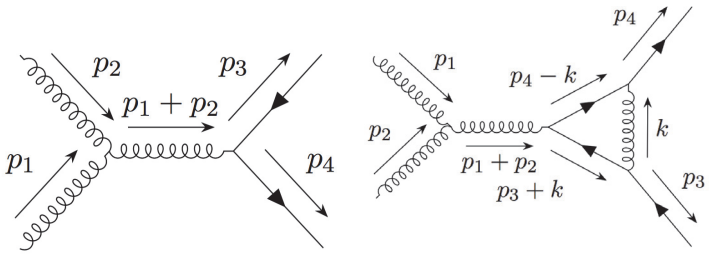
Figure 2.4: The Feynman rules of QCD with one quark flavor. The straight lines indicate quarks and the curly lines indicate gluons. External lines have the same properties as the corresponding fields and their properties (spin and color indices, momenta) are labelled with the corresponding number. The tensor V and W correlate the color and Lorentz structures of the incoming particles.

Because the vertices of Feynman are associated to factors of the interaction couplings, they can be used directly to find at which order a diagram contributes. Finding the perturbative expansion of a transition amplitude at a given order is therefore reduced to drawing all possible diagrams that have the right external legs and the right number of vertices. For a given number of external legs, higher orders in the expansion correspond to diagrams with more closed loops, and for each loop in the diagram, the momentum flowing through one line cannot be fixed by momentum conservation, as shown in Figure 2.5.

Part of the Feynman rules is the prescription that this unspecified momentum k must be integrated over using the d-dimensional measure

$$\int \frac{d^d k}{(2\pi)^d}. \tag{2.38}$$

One of the issues that arise when evaluating Feynman diagrams is that the integrals in the amplitudes associated to loop diagrams are often divergent in the limit $d \to 4$, and we need to regularize and renormalize the amplitudes. We will systematically work in the context of *dimensional regularization*, in which we use the number of dimensions as a regulator: we compute the integral for a generic $d = 4-2\epsilon$ and express it as a Laurent series in ϵ which we can analytically continue to complex ϵ.



(a) Tree-level diagram **(b)** Loop-level diagram

Figure 2.5: In tree-level diagrams, all internal momenta are fixed by the external legs while in loop diagrams, one internal propagator per loop is undetermined and integrated over.

Poles in the Laurent series of loop integrals have two origins: ultraviolet (UV) or infrared (IR). Ultraviolet poles come from divergences in the limit where the loop momentum k goes to infinity while infrared divergences potentially appear when one of the propagators in the loop integral goes to zero for some finite value of k.

Observables are made free of UV divergences in the limit $\epsilon \to 0$ through the process of renormalization, in which we relate the parameters that appear in the Lagrangian, called bare parameters, to finite, ϵ-independent quantities called the renormalized parameters. Let us consider for example the gauge coupling in QCD, g_s. The bare QCD coupling g_s^B appears in the term coupling quarks to gluons:

$$\mathcal{L}_{\text{QCD}} \supset g_s^B \bar{q} \gamma_\mu T^a q G_a^\mu. \tag{2.39}$$

Because S_{SM} is an action, it has no dimension in natural units where $\hbar = c = 1$, whence terms in the Lagrangian must be d-dimensional. The kinetic terms determine the dimension of the fields and we can use the interactions to fix the dimensionality of the coupling:

$$[q] = \frac{d-1}{2},\ [G^\mu] = \frac{d-2}{2},\ \left[g_s^B\right] = \frac{4-d}{2}. \tag{2.40}$$

If we want to define a finite renormalized coupling g_s^R independent of the number of dimensions, we need to compensate for the dimension of the bare parameter using an external scale μ which is a priori arbitrary. We will therefore define in general

$$g_s^B = \mu^\epsilon g_s^R(\mu) Z_{g_s}, \tag{2.41}$$

where Z_{g_s} is the ϵ and μ dependent *renormalization constant*. The renormalization constants are determined as an expansion in the renormalized coupling constant by imposing that all transition amplitudes in the theory be free of UV divergences in the limit $\epsilon \to 0$ once expressed in terms of the renormalized couplings. As a result, the renormalized constant depend on the precise way the divergences are absorbed. In this thesis, we will usually work in the modified minimal subtraction scheme ($\overline{MS}$ scheme), in which the relation between a bare parameter λ^B and its renormalized counterpart λ^R at a given order

in the perturbative expansion is

$$\lambda^B = S_\epsilon^{-\kappa_\lambda/2}\mu^{\kappa_\lambda\epsilon}\lambda^R\left(\sum_{n=1}^{N}\frac{Z_\lambda^{(n)}}{\epsilon^n}\right). \tag{2.42}$$

where $\left[\lambda^B\right] = 4 - \kappa_\lambda\epsilon$ and $S_\epsilon = (4\pi)^\epsilon \exp(-\gamma\epsilon)$, with γ the Euler-Mascheroni constant is a factor that systematically appears in loop integrals and is removed in $\overline{\text{MS}}$. In practice, we will drop the factors of S_ϵ for all parameters except for the strong coupling constant as a matter of simplicity. This yields differences of order $\mathcal{O}(\epsilon)$ in the finite results and therefore does not affect physical results.

The fact that absorbing divergences in a finite number of renormalization constants is enough to make all the transition amplitudes in a theory free of UV divergences is far from trivial and only happens for renormalizable theories such as the Standard Model [73].

Once the renormalization structure of a theory has been determined, renormalized amplitudes $\mathcal{A}^R$ can be related to the amplitude $\mathcal{A}$ expressed in terms of the bare parameters

$$\prod_\phi Z_\phi^{\frac{1}{2}}\mathcal{A}(\lambda^B) = \mathcal{A}^R(\lambda^R), \tag{2.43}$$

where the product is taken over the external fields and the Z_ϕ are the field renormalization constants.

After renormalization amplitudes can still contain infrared divergences. These cannot be absorbed by redefinitions of the parameters, but cancel against singularities that appear in the phase space integration if one considers IR-safe observables [74, 75], as we will discuss in Section 2.2.3.

2.2.2 From hadronic processes to partonic processes

Applying the techniques of perturbative QFT to quantum chromodynamics is a suitable approach to evaluating partonic cross sections for processes where the initial state is defined in terms of partons (quarks and gluons) Indeed, the renormalized strong coupling constant $\alpha_s^R(\mu) = \left(g_s^R(\mu)\right)^2/(4\pi)$ is a decreasing function of the renormalization scale as we showed in Figure 2.3 and it becomes smaller than 1 for $\mu \gtrsim 1\ \text{GeV}$. It is however impossible to perform parton scattering experiments directly because of confinement: we can only collide hadrons. High-energy hadron scattering can however be described in terms of elementary parton scattering thanks to factorization: due to the large separation between the typical hadronic energy scale ($\sim 1\ \text{GeV}$) and the scattering energy in experiments ($13\ \text{TeV}$ at the LHC), individual parton scattering probabilities are added incoherently up to power suppressed corrections that scale at worse like the inverse center-of-mass energy. The factorization formula for the proton-proton scattering cross section of the processus $pp \to X$ at a center-of-mass energy $\sqrt{s}$ takes the form

$$\sigma_{pp\to X} = \sum_{i_1,i_2}\int_0^1 dx_1 dx_2 f_{i_1}(x_1,\mu)f_{i_2}(x_2,\mu)\hat{\sigma}_{ij\to X}(\hat{s},\mu) + \mathcal{O}\left(\frac{1}{\sqrt{\hat{s}}}\right), \tag{2.44}$$

where the sum goes over the partons inside the protons, $\hat{\sigma}_{ij\to X}$ is the partonic cross section for the process $ij \to X$ at the *partonic* center-of-mass energy $\hat{s} = x_1x_2s$ and $f_i(x,\mu)$ is the parton distribution function (PDF) of the parton i in the proton, which can intuitively be interpreted as the probability of finding the parton i in the proton with a fraction x of its total momentum. The PDF depend on an

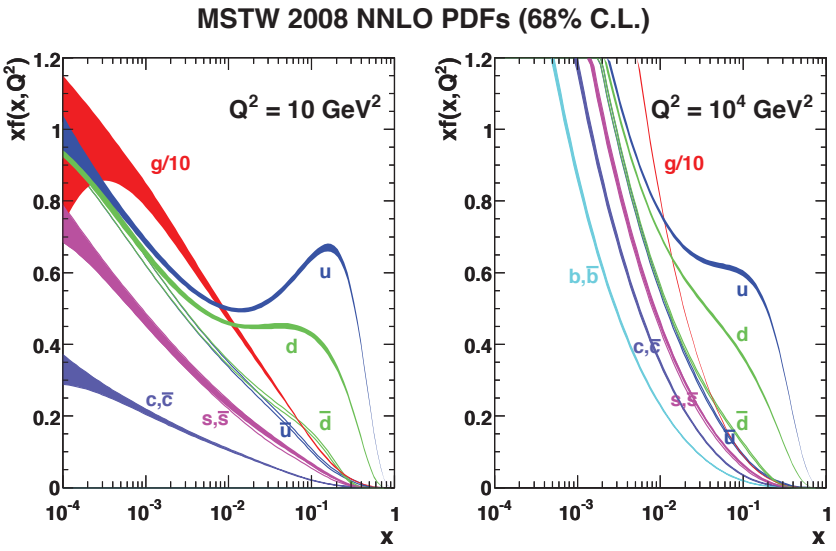


Figure 2.6: Examples of PDF from the MSTW2008 fit [77].

arbitrary scale called the factorization scale, which can be thought of as the scale that separates the hard physics described by $\hat{\sigma}$ and the low-energy physics described by the PDF. Factorization is only proven for a specific set of scattering processes but is conjectured to hold generally in high energy scattering [76].

Because the PDF relate to the internal dynamics of hadrons, they cannot be determined by perturbative QCD, and the best approach at the moment is to measure them experimentally through fits of experimental data. The result of such a fit is provided in Figure 2.6. The first historical measurements were performed using data from deep inelastic scattering (DIS) experiments in which electrons were collided with protons at high energies. Nowadays data is used from many experiments and many PDF fits are available [77–79].

Relating partonic final state to physical, hadronic final states seen in colliders is another part of QCD that cannot be predicted using perturbative QCD. At the energy regime of the LHC, the main experimental signatures of partonic final states with colored partons are hadronic jets, narrowly collimated bundles of hadrons. In electron-positron collisions at high energies for example, most hadronic final states are two-jet final states, in which two colimated sprays of boosted hadrons are back-to-back as shown in Figure 2.7, which corresponds to the leading-order partonic picture in which a pair of quarks is produced through a virtual photon or Z boson. From a theoretical point of view, the link from final states partons to hadronic states is described in two steps: fragmentation and hadronization. Fragmentation is still governed by perturbative processes and corresponds to the repeated splittings of partons ($g \to q\bar{q}$, $q \to qg$, ...). If the initial parton is very boosted, its fragmented version will consist of many partons going in the same directions, which will give rise to a jet. The perturbative picture breaks down when the typical momentum transfer in splittings becomes lower than $1\ \mathrm{GeV}$, at which point the description is switched to hadronization, which is the nonperturbative process in which partons group together to form physical hadrons. Hadronization models are not constructed from first principles, but are phenomenological models that are fitted to experiment.

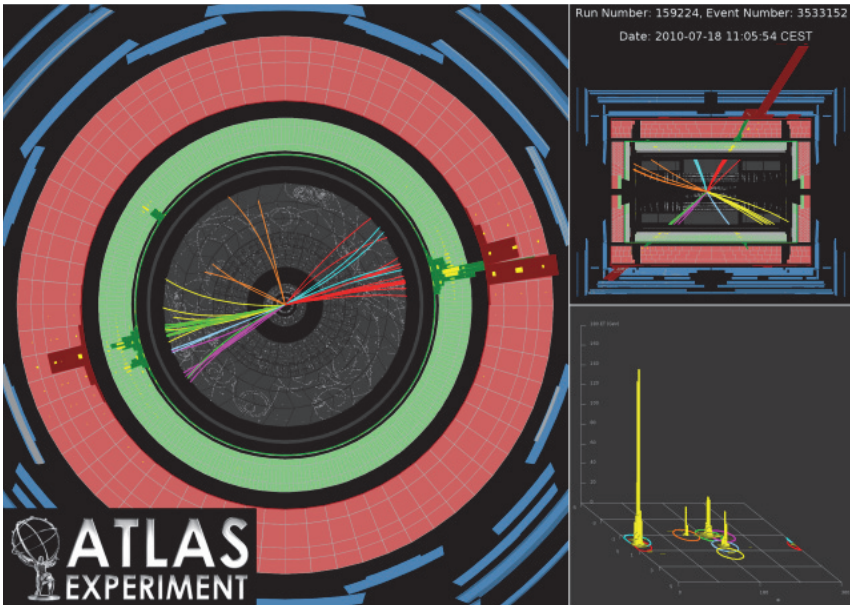


Figure 2.7: An example of a dijet event in the ATLAS detector at the LHC

2.2.3 Calculating partonic cross-sections at NLO

Let us briefly discuss the structure of NLO calculations in QCD at the parton level. As we explained in Section 2.2.1, higher order corrections to observables in QCD contain divergences that correspond either to UV divergences, or to IR divergences which disappear once we consider IR-safe observables, in which phase space integration divergences cancel against the IR divergences of loop integrals. We will illustrate our discussion with the cross-section for inclusive hadron production in electron-positron scattering well below the Z boson mass. The leading order contribution for this observable comes from the Feynman diagram shown in Figure 2.8. The total, unpolarized cross section σ_H is obtained by

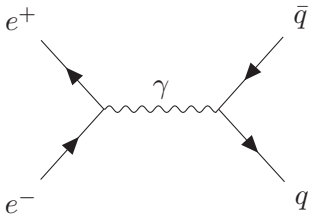
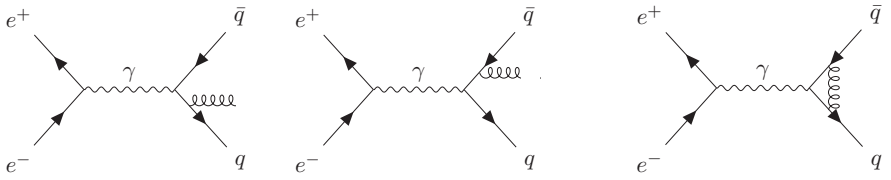


Figure 2.8: Leading-order contribution for $e^+e^- \to$ hadrons

integrating the norm of the amplitude over the phase space of the two quarks

$$\begin{aligned} \sigma_H &= \frac{4\pi\alpha^2}{s}\left(\sum_i \mathcal{Q}_i\right) &&+ \mathcal{O}(\alpha^2\alpha_s) \\ &= \sigma_0 &&+ \frac{\alpha_s}{2\pi}\sigma_1 + \ldots, \end{aligned} \tag{2.45}$$

where α is the electromagnetic coupling constant $\alpha = e^2/(4\pi)$, $\sqrt{s}$ is the center-of-mass energy of the collision, the $\mathcal{Q}_i$ are the quark charges and the sum runs over active quark flavors. The next-to-leading (NLO) corrections, which are $\mathcal{O}\left(\alpha^2\alpha_s\right)$ come from two different sources: virtual corrections, which include an extra loop, and real emission corrections, in which an extra parton is created, which will also hadronize. In the case of $e^+e^- \to$ hadron, they are realized in the Feynman diagrams shown in Figure 2.9. The $\mathcal{O}(\alpha_s)$ virtual correction σ_{1V} comes through the interference with the LO Feynman

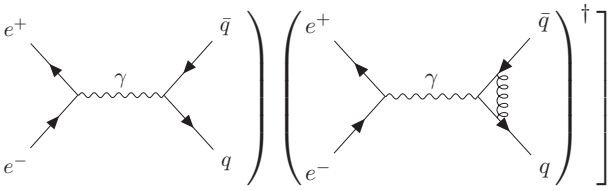


(a) Real emissions **(b)** Virtual correction

Figure 2.9: Feynman diagrams contributing to the NLO correction to the total cross section for $e^+e^- \to$ hadron

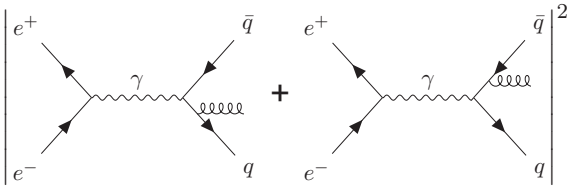
diagram, integrated over the phase space of the pair of quarks $d\Phi_2$, as shown in Equation 2.46.

$$\sigma_{1V} \propto \int d\Phi_2 \, 2\mathrm{Re} \left[\left(\quad \right) \left(\quad \right)^\dagger \right] \tag{2.46}$$



The contribution from real emissions on the other hand, does not interfere with the LO amplitude because the final state is different. Its contribution σ_{1R} is therefore directly the square of the real emission amplitude integrated over the phase space of the three-particle final state $d\Phi_3$, as shown in Equation 2.47.

$$\sigma_{1R} \propto \int d\Phi_3 \left| \quad + \quad \right|^2 \tag{2.47}$$



After renormalization, both contribution contain IR divergences coming respectively from the loop integral for the virtual correction and from the phase space integral for the real emission correction. Using dimensional regularization for both integrals we find that

$$\sigma_{1V} = \sigma_0 H(\epsilon) \left(-\frac{2}{\epsilon^2} - \frac{3}{\epsilon} - 8 + \mathcal{O}(\epsilon) \right) \tag{2.48}$$

$$\sigma_{1R} = \sigma_0 H(\epsilon) \left(\frac{2}{\epsilon^2} + \frac{3}{\epsilon} + \frac{19}{2} + \mathcal{O}(\epsilon) \right), \tag{2.49}$$

where

$$H(\epsilon) = \frac{3(1-\epsilon)^2}{(3-2\epsilon)\Gamma(2-2\epsilon)}. \tag{2.50}$$

It appears clearly that the divergences cancel in the total cross section:

$$\sigma_H = \sigma_0 + \frac{\alpha_s}{2\pi}\left(\sigma_{1R} + \sigma_{1V}\right) = \sigma_0\left(1 + \frac{\alpha_s}{\pi}\right). \tag{2.51}$$

It is a general feature of infrared divergences that those arising from real and virtual gluons must cancel if sufficiently inclusive observables are considered [74, 75]. Such observables are called infrared-safe observables and must verify the proprety that they measure the same value on a n-parton final state as on a $n+1$-parton final state in which the extra parton has a very small energy or is collinear with another parton. More explicitly, if S_n and S_{n+1} are the measure functions that define the observable as a function of the n and $n+1$ momenta of the final state,

$$\lim_{p_{n+1}\to 0} S_{n+1}(p_1, \ldots, p_n, p_{n+1}) = S_n(p_1, \ldots, p_n), \tag{2.52}$$

$$\lim_{p_{n+1}\to p_n} S_{n+1}(p_1, \ldots, p_n, p_{n+1}) = S_n(p_1, \ldots, p_n). \tag{2.53}$$

Clearly, inclusive hadron production is an infrared-safe observable, since the measure function is a constant independent of the number of partons. On the other hand, the cross section for exclusive $q\bar{q}$ production is not infrared safe, since the measure functions can be defined as $S_2(p_1, p_2) = 1$, $S_{n\neq 2}(\{p_i\}) = 0$. Indeed, if we try to evaluate it explicitly to NLO, we should perform the same calculation as before and drop the real emission diagram, hence spoiling the cancellation of divergences.

In general, the cross section for the inclusive production of any neutral particle is infrared-safe. For example, the cross section for Higgs production that we will consider in Chapter 6 is infrared safe because it is fully inclusive over the production of final state partons. It is still possible to define exclusive observables, which for example restrict the presence of hadronic jets coming from energetic partons, but even such observables cannot entirely restrict the presence of soft hadronic activity as jets must be defined in an infrared-safe way.

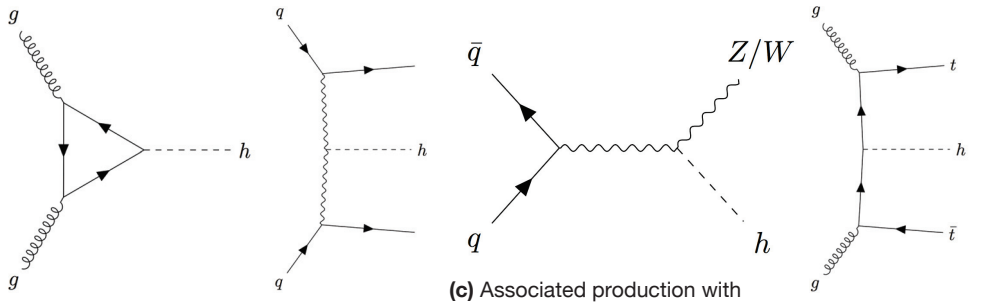
2.3 Higgs boson physics at the LHC

2.3.1 Current experimental status of the Higgs boson

Since the discovery of the Higgs boson in 2012 [5, 6], it has been extensively studied at the LHC to test its compatibility with the SM. After the Run I, a number of properties have been confirmed to be consistent with the SM. In particular, the experimental collaborations have confirmed that the particle is a CP-even state compatible with a spin-0 particle [80, 81]. The spin-1 hypothesis has been conclusively excluded and, while the hypothesis of a spin-2 is difficult to rule out entirely as many possible models exist for spin-2 production and decay [82], but all the models tested experimentally were excluded with a large confidence level while the spin-0 hypothesis is compatible with the experimental data.

A SM Higgs boson with a mass of $125\ \mathrm{GeV}$ has a rich phenomenology, as shown in Figure 2.11. The main production mechanism is gluon-fusion, which dominates the cross section by contributing more than 90%, followed by vector boson fusion and associated production with vector bosons. Rarer production modes include associated production with top and bottom quarks as well as the pair production of Higgs bosons. LO Feynman diagrams for each of these processes are show in Figure 2.10. It is in-

teresting to note that the main production channel is loop induced; its dominance is explained by the importance of gluon PDFs at high energies and the large coupling of the top quark to the Higgs boson.



(a) Gluon fusion **(b)** Vector boson fusion **(c)** Associated production with W/Z **(d)** Associated production with $t\bar{t}$

Figure 2.10: Sample Feynman diagram for each of the main channels for Higgs boson production.

The decays of the Higgs boson are also diverse, as already three decay channels (ZZ, WW, $\gamma\gamma$) contributed to the initial discovery [5, 6]. Several of the dominant decay channels are however difficult to separate from the background and have yet to be conclusively observed, including the decay to bottom quark pairs, which is the most frequent decay mode.

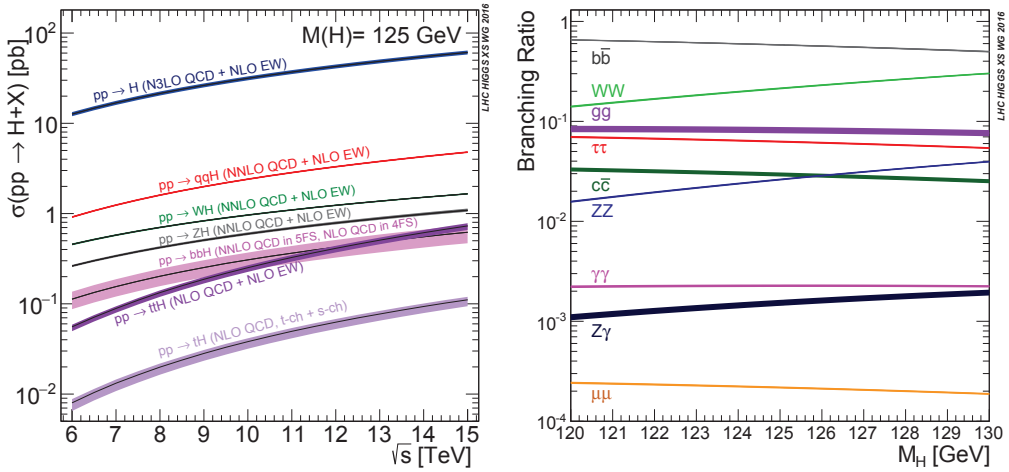


Figure 2.11: Production and decay modes for a $125\ \mathrm{GeV}$ Higgs boson at the LHC from the CERN Yellow report 4 [83].

A comprehensive picture of Higgs boson interactions has already been established by the experimental collaborations using the complete dataset of Run I [7]. The approach taken in this combined study was to perform fits on parametrizations of the full dataset, using three different parametrizations. The first version is very agnostic and uses products of the form $\sigma_i \times B_j$ as fit parameters, where σ_i is the production cross section for some production mode and B_j is the branching ratio for some decay. For each analysis A, the observed number of event n_A is compared with a theoretical prediction of the form

$$n_A = \mathcal{L} \sum_{ij} \epsilon_{ij}^A \left(\sigma_i \times B_j\right) \frac{(\sigma \times B)_{\text{ref}}}{(\sigma \times B)_{\text{ref}}^{\text{SM}}}, \tag{2.54}$$

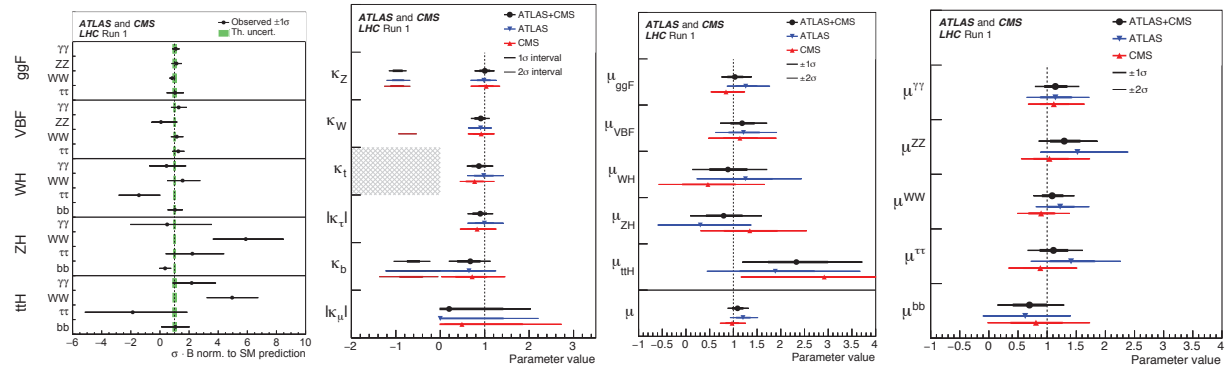
where $\mathcal{L}$ is the integrated luminosity and ϵ_{ij}^{A} is the selection efficiency of the channel ij in the analysis A, which is determined by simulations. The processus $gg \rightarrow H \rightarrow ZZ$ is used as a reference for the overall normalization due to its overall reliability, being a channel with low backgrounds and small systematic uncertainties. The result of this fit is shown in Figure 2.12a. The second parametrization is done in terms of signal strengths, which works under the assumptions that the measurements of production and decay strengths can be factorized universally. The fit parameters μ_i for production and μ_j for decays rescale the theoretical predictions for the contribution of each mode to the measured yield of each analysis:

$$n_A = \mathcal{L} \sum_{ij} \epsilon_{ij}^{A} \, \mu_i \mu_j \, \sigma_i^{\mathrm{SM}} \times B_j^{\mathrm{SM}}. \tag{2.55}$$

The result of this fit is shown in Figure 2.12c. The third parametrization is based on the idea of *coupling modifiers* often referred to as the κ-*framework*. A coupling modifier κ_i is associated to each coupling involving the Higgs boson and the contribution of each channel is rescaled by a combination kappas based on the LO scaling of the cross section times branching ratio. For example, in search for the associated production of a Higgs boson with a pair of top quarks, with the Higgs boson decaying into bottom quark, the signal contribution (*i.e.* the contribution from genuine $t\bar{t}H$, $H \rightarrow b\bar{b}$) contributes to the event number n in the following way

$$n = \mathcal{L}\epsilon \, \kappa_t \kappa_b \, \sigma_{t\bar{t}H}^{\mathrm{SM}} \times B_{H\rightarrow b\bar{b}}^{\mathrm{SM}} + \ldots \tag{2.56}$$

where κ_t (κ_b) is the coupling modifier for the top (bottom) Yukawa coupling. The result of this fit is shown in Figure 2.12b. This approach has a nice intuitive interpretation since, at leading order, chang-



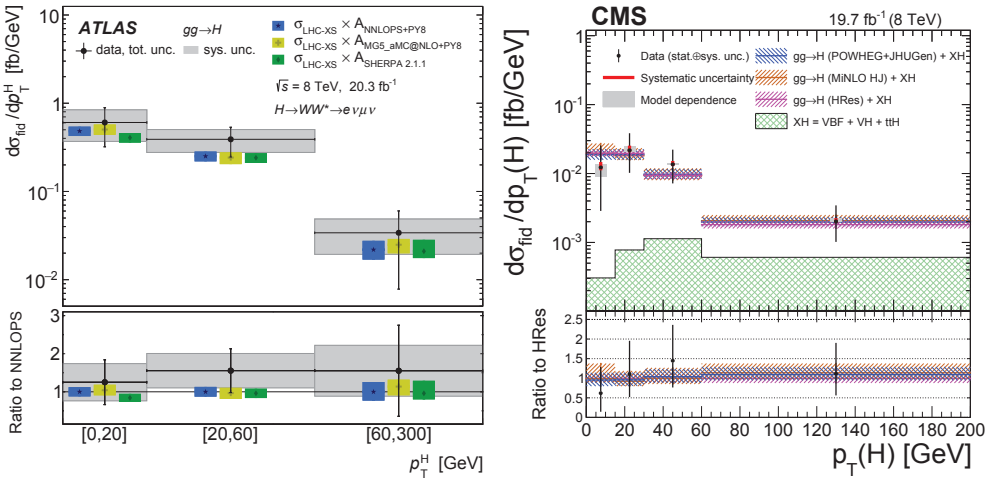
(a) $\sigma \times B$ parametrization **(b)** κ parametrization **(c)** signal strength parametrization

Figure 2.12: Results for the fits in the three parametrizations used in [7].

ing the coupling of the Higgs boson to a given SM particle is equivalent to finding a value of the corresponding kappa different from one. However, predictions in the Higgs sector are very sensitive to radiative corrections so that that the relationship between a modified interaction strength and a measured kappa can be nontrivial. Furthermore, the phenomenology of SM extensions can be different to that of the SM. As a result, relating the measurement of kappas to the interactions in a specific BSM model can be non-trivial. One should therefore be wary of interpreting the kappas as BSM coupling

modifiers as they are really fit parameters that cannot be trivially related to the parameters of some QFT.

The bulk of the Run I Higgs analysis campaign was centered around total rates, which allowed to answer the burning questions that followed the initial discovery: is the Higgs boson produced as we expect? Does it decay in the right channels? As more data is being gathered and analyzed, we have access to a growing body of information not only on the rates, but also on the kinematics of Higgs-related processes. An important milestone in this regard was the publication of differential rates for Higgs boson production [84, 85], which provided transverse momentum spectra, as shown in Figure 2.13. Such measurements will provide crucial tools to test the SM down to the smallest details as they can



(a) ATLAS $(gg \to H \to WW)$ **(b)** CMS $(gg \to H \to ZZ)$

Figure 2.13: Transverse momentum spectra for the Higgs boson after Run I.

shed light on the precise nature of the Higgs boson couplings, and in particular probe the presence of new Lorentz structures in the interactions.

2.3.2 Calculating Higgs properties in the SM and the HEFT

Gluon fusion and the infinite top mass limit

Testing the properties of the Higgs boson at colliders implies having reliable theoretical predictions for its properties. An important effort has been expanded to provide a comprehensive body of theoretical work to support the work of experimenters in studying Higgs physics at the LHC [34, 35, 40, 44, 45, 50], and as shown in Figure 2.11, the dominant modes are known to high orders in QCD.

In this section, we will focus on the specifics of the main production mode, gluon fusion, which is a loop-induced process: the LO Feynman diagrams for this mode include a loop of top quarks, as shown in Figure 2.14. The amplitude associated to these Feynman diagrams can be evaluated easily and takes

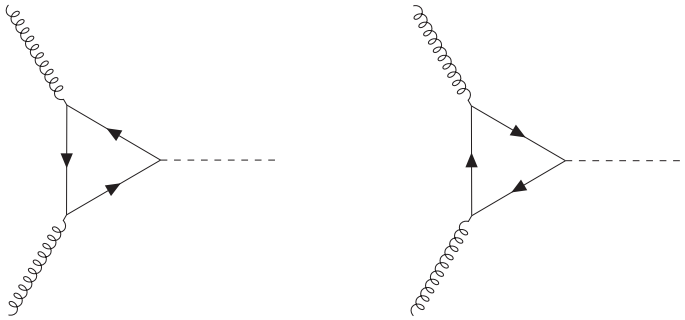


Figure 2.14: LO diagrams for Higgs production through gluon fusion. The dominant contribution comes from top quarks running in the loop.

the form

$$-i\frac{\alpha_s}{2\pi v}\delta^{ab}\varepsilon_\mu(p_1)\varepsilon_\nu(p_2)\left(p_1\cdot p_2 g^{\mu\nu}-p_1^\nu p_2^\nu\right)\tau\left(1+(1-\tau)\tan^{-1}\left(\frac{1}{\sqrt{\tau-1}}\right)^2\right), \tag{2.57}$$

where $\tau = \frac{4m_t^2}{m_h^2}$, the $\varepsilon^\alpha(p_i)$ are the polarization vectors of the gluons and the indices a, b their color indices. Radiative corrections to this process are notoriously difficult to obtain: the NLO correction was calculated in 2006 [51, 52] and further exact results are still at a bottleneck due to the complexity of evaluating multiloop integrals with internal masses. A successful alternative approach has been to observe that despite the fact that the mass of the top quark is not much larger than the mass of the Higgs boson, expanding the amplitude in the limit $m_t \to \infty$ is a reasonable approach due to the fact that the expansion is formulated in powers of $1/\tau \simeq 0.13$. In this limit, the LO amplitude reduces to

$$-i\frac{\alpha_s}{3\pi v}\delta^{ab}\varepsilon_\mu(p_1)\varepsilon_\nu(p_2)\left(p_1\cdot p_2 g^{\mu\nu}-p_1^\nu p_2^\nu\right), \tag{2.58}$$

which is independent of the top mass and corresponds to a local interaction between the Higgs boson and gluons. Indeed, the same amplitude can be obtained from an effective interaction Lagrangian

$$\mathcal{L}_{\text{eff}} = -\frac{1}{4}\frac{\alpha_s}{3\pi v}hG_a^{\mu\nu}G^a_{\mu\nu}, \tag{2.59}$$

which generates a tree level Higgs-gluon interaction, as shown in Figure 2.15.

The Higgs Effective Field Theory

A systematic approach can be taken to improve the prediction for gluon fusion in the large top mass limit using the framework of effective field theory (EFT) [86]. Indeed, taking this limit is tantamount to considering an EFT in which one integrates the top quark field out of the theory [53–55]. This procedure generates non-renormalizable interactions between Standard Model fields and modifies the values of the couplings, all of which are suppressed by powers of $1/m_t$. Performing this task in general would be a very involved operation since removing the top quark from the theory breaks the gauge invariance. On the other hand, if one is only interested in evaluating the QCD corrections to gluon fusion, it is

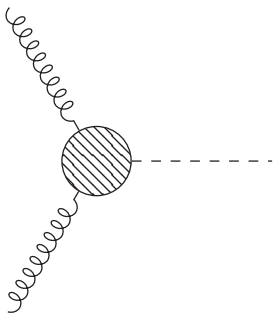


Figure 2.15: Effective coupling between the Higgs boson and the gluons in the infinite top quark limit.

sufficient to consider a "limited SM" containing QCD and the Higgs boson:

$$\begin{aligned}\mathcal{L}_{\text{QCD+h}} = &-\frac{1}{4}G_a^{\mu\nu}G^a_{\mu\nu} + \sum_{f\neq t} i\bar{q}^f \not{D} q^f + \bar{t}\left(i\not{D} - m_t\right) t \\ &+ \frac{1}{2}\left(\partial_\mu h \partial^\mu h - m_h^2 h^2\right) - \sum_f \frac{y_f}{\sqrt{2}}\bar{q}qh + V(h).\end{aligned} \tag{2.60}$$

Since the goal of the large top mass expansion is to describe gluon fusion, happening for partonic center-of-mass energies $\sqrt{\hat{s}} > 125\ \text{GeV}$, it is reasonable to consider all light flavors massless. Integrating out the top quark, we can write a generic form for the new theory, which is called the Higgs Effective Field Theory (HEFT):

$$\begin{aligned}\mathcal{L}_{\text{HEFT}} = &-\frac{1}{4}G_a^{\mu\nu}G^a_{\mu\nu} + \sum_f \bar{q}^f \left(i\not{D} - m_f^{\text{eff}}\right) q^f \\ &+ \frac{1}{2}\left(\partial_\mu h \partial^\mu h - \left(m_h^{\text{eff}}\right)^2 h^2\right) - \sum_f \frac{y_f^{\text{eff}}}{\sqrt{2}}\bar{q}qh + V^{\text{eff}}(h) \\ &+ \frac{1}{m_t}\mathcal{L}^{(5)}_{\text{eff}} + \mathcal{O}\left(\frac{1}{m_t^2}\right)\end{aligned} \tag{2.61}$$

where $\mathcal{L}^{(5)}_{\text{eff}}$ contains non-renormalizable dimension five operators and the ellipsis contains a tower of higher dimensional operators suppressed by powers of $1/m_t$. We need to consider all the parameters appearing in this Lagrangian as new parameters, called Wilson coefficients, which we need to match to the parameters of the complete theory by comparing their prediction for some choice of observables in the limit $m_t \to \infty$. This comparison will in general yield an expression for the parameters of the HEFT in the form of an expansion

$$\lambda^{\text{eff}} = \lambda_0 + \frac{\lambda_1}{m_t} + \frac{\lambda_2}{m_t^2} + \ldots, \tag{2.62}$$

where the λ_i are expressed in terms of the parameters of the full theory. It is of course true that the couplings of all operators that are present both in the full theory and in the HEFT should match at the $\mathcal{O}\left(1/m_t^0\right)$ level. For example, let us perform the leading-order matching of a light-quark Yukawa

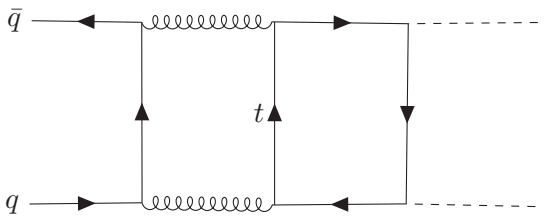


Figure 2.16: The lowest order SM diagram that will yield a contribution to the Wilson coefficient of the operator $\bar{q}qhh$ in the HEFT.

coupling using the amplitudes for the process $H \to q^f \bar{q}^f$ in the SM, $\mathcal{A}$, and in the HEFT, $\mathcal{A}^{\text{eff}}$:

$$\mathcal{A} = -i\frac{y_f}{\sqrt{2}}\bar{u}v, \tag{2.63}$$

$$\mathcal{A}^{\text{eff}} = -i\frac{y_f^{\text{eff}}}{\sqrt{2}}\bar{u}v + \mathcal{O}\left(\frac{1}{m_t}\right). \tag{2.64}$$

By imposing that the expressions in Equations 2.63 and 2.64 match up to power-suppressed terms, we find that $y_f^{\text{eff}} = y_f + \mathcal{O}(1/m_t)$. The operators in $\mathcal{L}_{\text{eff}}^{(5)}$, on the other hand must describe top-mediated interactions that cannot be reduced to operators already present in the SM. We have already found one of them above and have performed the matching to the full SM at leading order:

$$\frac{1}{m_t}\mathcal{L}_{\text{eff}}^{(5)} \supset -\frac{C}{4}hG_a^{\mu\nu}G_{\mu\nu}^a \tag{2.65}$$

$$C = \frac{\alpha_s}{3\pi v}. \tag{2.66}$$

Of course, other operators must be considered in order to describe the theory completely. For example, the operator $\bar{q}qhh$ can be matched to the SM using the process $\bar{q}q \to hh$. As shown in Figure 2.16, the first Feynman diagrams that contains power-suppressed terms in the limit $m_t \to \infty$ appears at the two loop levels. As a result, the Wilson coefficient of this operator will be of order $\mathcal{O}\left(\alpha_s^2\right)$.

Radiative corrections in the HEFT

Despite the fact that the HEFT is non-renormalizable, it verifies the property of *order-by-order renormalizability*, which is a generic feature of EFTs. Explicitly, if one truncates the HEFT Lagrangian at a given order n in the $1/m_t$ expansion, the set of operators present in the theory is sufficient to absorb all the divergences that appear in observables if they are themselves truncated at the order n. In particular, this means that one can safely consider QCD corrections of HEFT observables. For example, the one-loop QCD correction to gluon fusion in the HEFT contains UV divergences that can be absorbed by dimension 5 counterterms, as shown in Figure 2.17.

The constraint that observables should be truncated at a given order in the expansion takes a very simple form for the first order in the expansion: one should only consider diagrams with at most one dimension 5 operator. For example, the Feynman diagram shown in Figure 2.18 cannot be renormalized in the HEFT truncated at dimension 5 because it contains two dimension 5 operators, making a

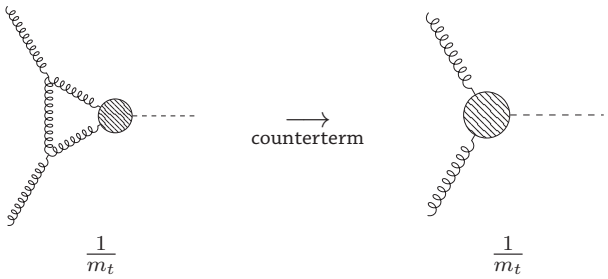
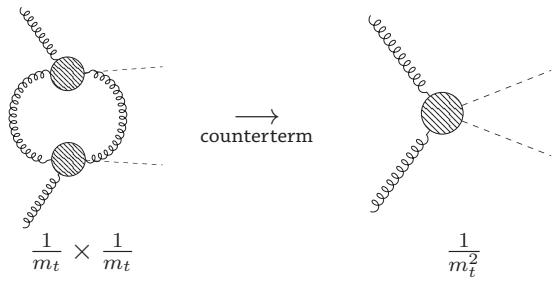


Figure 2.17: Example of a Feynman diagram involving a dimension 5 operator whose divergence can be absorbed by a dimension 5 operator. This is possible because this Feynman diagram contributes to the amplitude at the same order in $1/m_t$ as the truncation of the HEFT.

dimension 6 counterterm necessary.



(2.67)

Figure 2.18: Example of a Feynman diagram involving dimension 5 operators whose divergence cannot be absorbed by a dimension 5 operator. Here, the contribution of the diagram to the amplitude is $\mathcal{O}\left(1/m_t^2\right)$, so a dimension 6 counterterm is necessary.

If we perform the truncation appropriately, it is possible to evaluate radiative corrections to the gluon fusion process without having to consider the massive quark loops that are still out of reach of exact calculations. Two types of corrections have to be considered for such calculations: radiative corrections to HEFT diagrams and corrections to the Wilson coefficients. Indeed, the Wilson coefficients are expressed as a double expansion in the coupling constants and the top mass, *i.e.* the λ_i in Equation 2.62 are themselves expressed as a series in α_s, which can be obtained by matching radiative corrections to observables between the HEFT and the SM. In the case of the gluon-fusion operator, this matching has been performed to high orders [87, 88] and for example at NLO, the renormalized Wilson coefficient gets a correction:

$$C = \frac{\alpha_s}{3\pi v}\left(1 + \frac{\alpha_s}{2\pi}\frac{11}{2}\right). \tag{2.68}$$

Calculations performed in the HEFT have been essential in providing trustworthy predictions for the gluon fusion cross section. Indeed, the NLO correction in the complete SM change the LO result by 70% at $\sqrt{s} = 8\ \mathrm{TeV}$, motivating the need to evaluate NNLO corrections, which were computed in the HEFT in 2002 [89]. This result yielded another 20% increase in the total cross section, showing that the expansion was converging, though slowly. The N^{3}LO corrections were obtained in 2015 [40] in an expansion around the threshold limit $s \to m_h^2$, yielding this time a correction to the prediction smaller than 5%. The slow convergence of the perturbative expansion for the gluon fusion cross section

is shown in Figure 2.19.

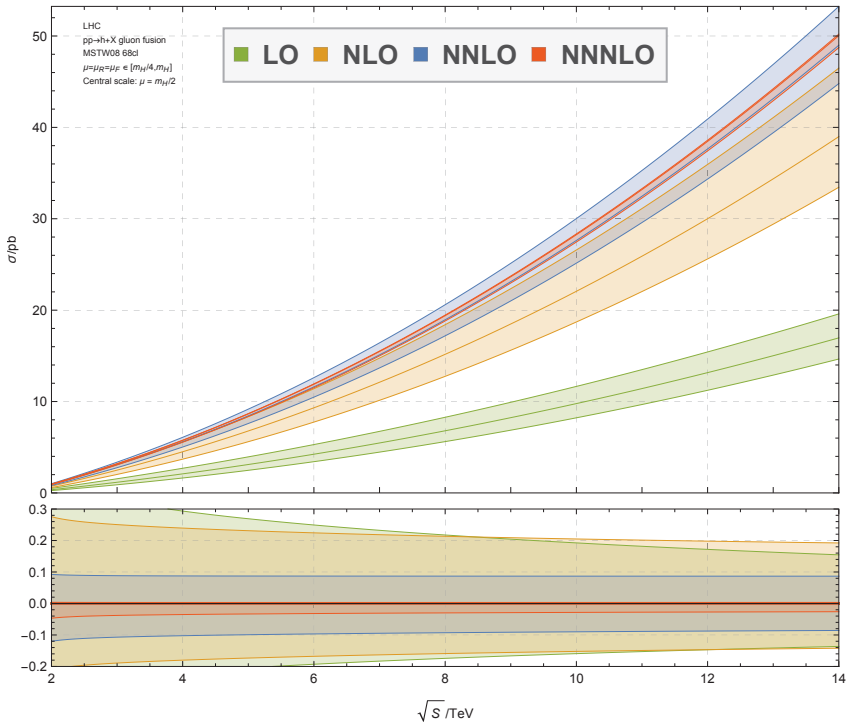


Figure 2.19: Gluon fusion cross section in pp collisions in the HEFT as a function of the center-of-mass energy at four different orders in perturbation theory. The bands around the central values reflect the scale dependence of each prediction.

2.4 The standard model effective field theory

2.4.1 Model independent tests of the Standard Model

Two main options exist to test for the existence of new physics at the LHC: direct searches and precision measurements. Direct searches test the hypothesis that some new particles are light enough to be produced at the LHC and aim at observing the signature of their production experimentally. The typology of possible signature is extremely varied and can range from clear-cut resonances in invariant-mass distributions [90–93] in the case of simple resonances, to more subtle signals with missing energy as considered in searches for dark matter [94], or very complicated final states characteristic of R-parity conserving decays of supersymmetric particles [95, 96]. Direct searches face the problem that many BSM models exist, and many of the most popular ones such as the minimal supersymmetric standard model (MSSM) [56, 57] have a very large parameter space and a correspondly rich and varied phenomenology [97]. Other types of approaches to BSM physics have many different realizations such as extra dimensions [98–102] or composite Higgs models [58, 59], making it difficult to devise methods to broadly constrain them. Because experimental searches target signatures and not models, a very popular approach in modern analyses is to use simplified models [103], which specify minimal particle spectra to probe specific signatures, which can then be reinterpreted in terms of complete models and be included in global fits [104]. Searches in the context of simplified models are therefore a good way

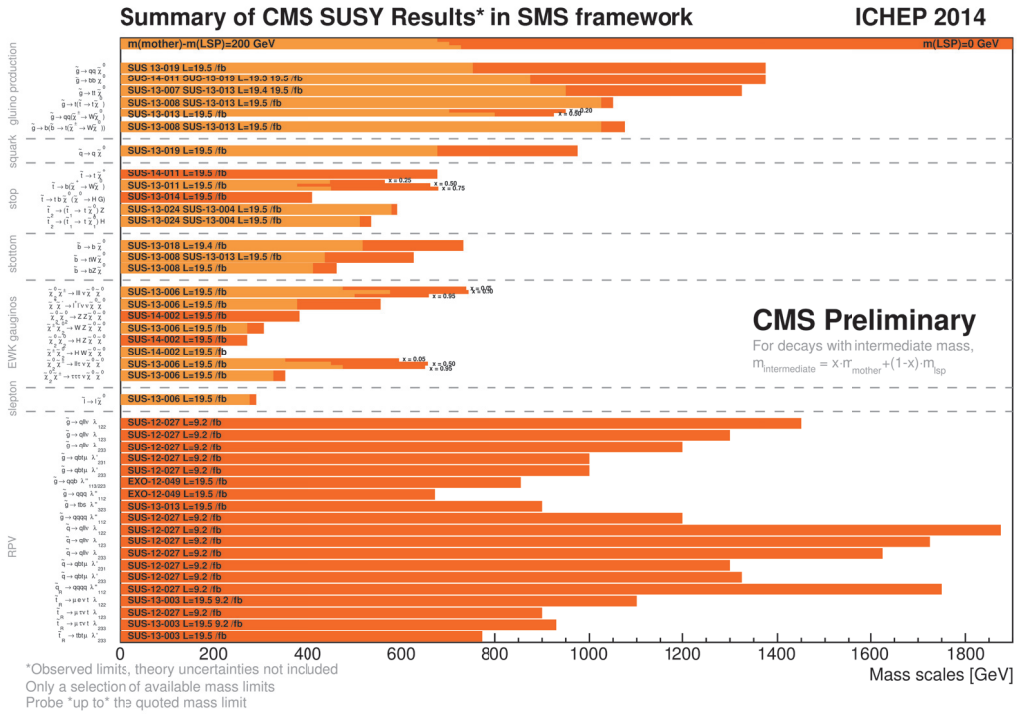


Figure 2.20: Results summary for supersymmetry searches from CMS using simplified models.

to make model independent statements on specific signatures, and can provide clear-cut limits such as those shown in Figure 2.20. Note however that the limits from simplified models usually need reinterpretation, as they often make simplifying assumptions such as the existence of a single decay channel for a given particle.

As no direct search for BSM physics has returned a positive result since the start of the LHC, a complementary type of analysis is gaining momentum, based on precision measurements. The idea behind this approach is that new physics particles might be too heavy to be produced at the LHC and could therefore be considered in the decoupling limit [105] as inducing new interactions between SM particles, much like weak decays could be described by an effective four-fermion interactions at an energy scale well below the W boson mass. From a BSM model $\mathcal{M}$ with only heavy new particles, one can therefore construct an effective theory [60–62] that describes the effects of new physics as corrections to SM observables in an expansion in the inverse power of the heavy mass scale Λ of the new particles after they have been integrated out of the theory. The Lagrangian for this effective theory is then expressed as an expansion in this heavy scale:

$$\mathcal{L}_{\text{EFT}}^{\mathcal{M}} = \mathcal{L}_{\text{SM}} + \sum_{i=1}^{\infty} \sum_{j=0}^{N_i} \frac{c_{j(\mathcal{M})}^{(i)}}{\Lambda^i} O_j^{(i)}, \tag{2.69}$$

where $\mathcal{L}_{\text{SM}}$ is the SM Lagrangian, the operators $O_j^{(i)}$ have an energy dimension of $4+i$ and the $c_j^{(i)(\mathcal{M})}$ are dimensionless couplings that are expressed in terms of the parameters of the BSM model $\mathcal{M}$, and therefore often verify relations among themselves that reflect the structure and symmetries of $\mathcal{M}$. An

essential feature of the effective theories obtained by decoupling heavy BSM particles is that the set of operators that can appear in this expansion at each order is finite[3]. As a result, it is possible to construct a general effective field theory, the standard model effective field theory (SMEFT), that can describe the decoupling limit of any BSM model. The Lagrangian of the SMEFT can therefore be expressed as

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_{i=1}^{\infty} \sum_{j=0}^{N_i} \frac{c_j^{(i)}}{\Lambda^i} O_j^{(i)}, \tag{2.70}$$

where this time the sum over operators covers all possible operators that verify the symmetries of the SM and the c_i are all independent couplings. In general, an observable in the SMEFT $\mathcal{O}$ is expressed as a series which matches the SM prediction $\mathcal{O}_{\text{SM}}$ in the limit $\Lambda \to \infty$

$$\langle \mathcal{O} \rangle = \langle \mathcal{O}_{\text{SM}} \rangle + \frac{1}{\Lambda} \langle \mathcal{O}_1 \rangle + \frac{1}{\Lambda^2} \langle \mathcal{O}_2 \rangle + \dots \tag{2.71}$$

where each order in the expansion depends on the Wilson coefficients $c_j^{(i)}$. A sufficient criterium for the applicability of the EFT approach for a given BSM model is that when one matches the parameters of the SMEFT with those of the model, the expansion of observables be convergent for LHC energies. From a bottom up point of view, the parameter space of the SMEFT considered in LHC studied should be restricted to the region where the expansion of the observables are convergent.

One can perform model independent tests of the SMEFT by fitting its prediction to experimental results for SM observables at the LHC [107–109]. In the case of a measurement compatible with the SM prediction, bounds can be placed on the SMEFT parameter space, while a small deviation can be fitted to the SMEFT prediction to measure the value of the SMEFT Wilson coefficients. The strength of the SMEFT in this procedure is its universality: by matching decoupling BSM models to the SMEFT, one can establish a correspondance between the Wilson coefficients and the model parameters [110], hence propagating constraints or measurements to a large number of BSM models.

2.4.2 The SMEFT at dimension six

Most current studies in the SMEFT are performed in a truncation at the second order in the expansion. Indeed, it would be insufficient to truncate at the first order, as the dimension 5 operators in the SMEFT are very restricted: up to the flavor structure, there is a single operator, the Weinberg operator [60]:

$$\mathcal{L}_5 = \frac{c_{\alpha\beta}}{\Lambda} \left((\bar{L}_\alpha \phi^c)(\phi^{c\dagger} L_\beta) + \text{h.c.} \right), \tag{2.72}$$

where α, β are flavor indices. This operator can be related to the existence of neutrino masses and other lepton-number-violating effects. Due to the very small value of the neutrino masses, this operator is heavily constrained and is expected to bear no impact on LHC physics. Indeed, these constraints imply $\frac{\Lambda}{c_{\alpha\beta}} \gtrsim 10^{14}$ GeV.

The situation changes dramatically when one considers the SMEFT expansion up to dimension six. Indeed, there are 59 independent operators that preserve baryon and lepton numbers and verify flavor

[3] A systematic approach to finding the number of operators at each order has been proposed in [106].

universality [61, 111], which yields a total of 76 real parameters, accounting for possible CP violations, and 2499 operators if a general flavor structure is considered [112]. This large number of operator ensures that essentially all areas of LHC physics can be modified by SMEFT effects, providing a good starting point to put model independent constraints on heavy BSM physics. The number of operators by itself is not a trivial result as one can find many relations between operators. The first really minimal basis was the Warsaw basis [111], and other bases exist such as the SILH basis [113].

If one constrains the SMEFT Lagrangian to conserve lepton number, the SMEFT truncated at dimension six contains does not contain dimension five operators

$$\mathcal{L}^{(6)}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \frac{1}{\Lambda^2}\sum_i c_i O_i, \tag{2.73}$$

where we drop the (2) superscript because we work at a fixed order. As a result, any observable $\mathcal{O}$ should be truncated at the order $\mathcal{O}\big(1/\Lambda^2\big)$ and contain no $1/\Lambda$ term, which makes it take a simple form:

$$\langle \mathcal{O} \rangle = \langle \mathcal{O}_{\text{SM}} \rangle + \frac{1}{\Lambda^2}\sum_i c_i \langle \mathcal{O}_i \rangle , \tag{2.74}$$

where $\langle \mathcal{O}_i \rangle$ is the result of the interference of the Standard Model with the EFT:

$$\langle \mathcal{O}_i \rangle = \sum_p \int d\,\Phi 2\text{Re}\left(\mathcal{A}_{\text{SM}}(p)\mathcal{A}_i^\dagger(p)\right)\mathcal{O}, \tag{2.75}$$

where the integration is taken over the phase space, the sum is taken over the relevant scattering processes, $\mathcal{A}_{\text{SM}}(p)$ is the SM amplitude for each process and $\mathcal{A}_i(p)$ is the amplitude for each process with one insertion of the operator O_i.

As in the case of the HEFT, the SMEFT is renormalizable order by order, and the fact that all operators appear at most linearly in observables makes the renormalization of the Wilson coefficients take a simple form:

$$c_i^B = Z_{ij}c_j^R, \tag{2.76}$$

where Z is a renormalization matrix which can mix operators during renormalization. The presence of dimension six operators also brings corrections to the SM parameters, for example the top mass has an expansion

$$m_t^B = m_t^R + \delta m_{\text{SM}} + \frac{1}{\Lambda^2}\sum_i c_i \delta m_t^{(i)}. \tag{2.77}$$

Note that the linear structure of the renormalization is a specific property of the truncation at dimension six. If we were to consider dimension eight operators, we would instead find renormalization relations of the form

$$c_i^{(4)B} = X_{ij}c_j^{(4)} + Y_{ijk}c_j^{(2)}c_k^{(2)}. \tag{2.78}$$

The general one-loop renormalization structure of the SMEFT has been determined in a series of three papers [112, 114, 115] that calculated the complete renormalization matrix of the SMEFT and the RGE dependence of the Wilson coefficients. Based on this result, a number of SMEFT predictions have been calculated at the LHC including radiative QCD corrections [116–124].

3 Multi-loop techniques

3.1 Feynman diagrams, topologies and master integrals

3.1.1 Dealing with many diagrams and integrals

The basic building blocks of perturbative calculations in Quantum Field Theory are encoded in Feynman diagrams, which provide a diagrammatic description of the terms that contribute to the calculation of transition amplitudes at a given order. Feynman diagrams themselves are constructed from the *Feynman rules* of the theory, which describe how elementary interactions between fields contribute to physical processes.

In a given theory, writing down the perturbative expansion for a given scattering amplitude is reduced to finding the set of diagrams that contribute to the desired order, *i.e.* with the correct number of vertices. While the discovery of Feynman diagrams has tremendously reduced the complexity of calculations in perturbative quantum field theory, modern calculations make finding the relevant set of Feynman diagrams for a given calculation a tremendous task. Indeed, the number of diagrams for processes with many loops and legs can reach the hundreds or thousands, making the use of computer programs to generate and evaluate Feynman diagrams essential. Several public codes exist that allow users to specify the Feynman rules of a theory and produce either abstract representation of the diagrams [125], or directly the expression for the corresponding terms in the transition amplitude [126].

One of the most popular tools available is the program `QGRAF` [125], which is very efficient and flexible. We developed the program `qgraf-xml-parser` detailed in Appendix B to express the diagrams generated by `QGRAF` as input files for `FORM` [127] to perform the tensor algebra calculations and provide expressions suitable for `Mathematica`.

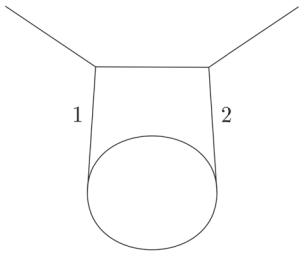


Figure 3.1: Example of a Feynman diagram where two propagators are identical in ϕ^4 theory. Indeed, the propagators labelled 1 and 2 are identical.

In general, a Feynman diagram $\mathcal{A}$ with n external legs takes the form

$$\mathcal{A} = \psi_1(p_1)^{\mathbf{a_1}} \dots \psi_n(p_n)^{\mathbf{a_n}} T_{\mathbf{a_1} \dots \mathbf{a_n}}, \tag{3.1}$$

where the ψ_i are the wavefunctions associated to each external leg and have indices, which we grouped into $\mathbf{a_i}$, describing their transformation properties under Lorentz and gauge transformations. It is in general much easier to work with scalar amplitudes when one works at loop level, as scalar integrals are in general easier to handle than tensor integrals. As a result, it is often useful to decompose the amputated amplitude T on a basis of tensor structure $\{B_i^{\mathbf{a_1} \dots \mathbf{a_n}}\}_{1 \leq i \leq N}$:

$$T_{\mathbf{a_1} \dots \mathbf{a_n}} = T_1 B_1^{\mathbf{a_1} \dots \mathbf{a_n}} + \dots + T_N B_N^{\mathbf{a_1} \dots \mathbf{a_n}}. \tag{3.2}$$

This basis only depends on the external legs, meaning that the decomposition of T is the same for all diagrams contributing for a given amplitude at all loop orders. For example, the amplitude for Higgs production through gluon fusion $\mathcal{A}(gg \to H)$ can be decomposed in a single tensor structure:

$$\mathcal{A}(gg \to H) = \varepsilon^{\mu_1} \varepsilon^{\mu_2} T^{a_1 a_2}_{\mu_1 \mu_2} = \varepsilon^{\mu_1} \varepsilon^{\mu_2} \delta^{a_1 a_2} \left(p_1 \cdot p_2 g_{\mu_1 \mu_2} - p_1^{\mu_1} p_2^{\mu_2} - p_1^{\mu_2} p_2^{\mu_1} \right) T_1, \tag{3.3}$$

where the ε^{μ_i} are the polarization vectors of the gluons and the a_i are the color indices.

Let us consider one scalar form factor T among the T_i in an L-loop amplitude evaluated in d dimensions. It takes the general form

$$T = \sum_j A_j \left(\{g_m\}, \{s_n\} \right) \int \prod_{l=1}^{L} \frac{d^d k_l}{(2\pi)^d} \frac{\mathcal{N}_j}{D_1^{n_1} \dots D_N^{n_N}}, \tag{3.4}$$

where the A_j are rational functions that depend on the couplings g_m and the external scales s_n, the numerators $\mathcal{N}_j$ are products of contractions involving at least one loop momentum and the D_i are propagator denominators of the form $\left(q^2 - m^2\right)$ where q is some combination of momenta and m some mass, and the $n_i \in \mathbb{N}^+$. Note that in most cases, all $n_1 = \dots = n_N = 1$, but in some cases, several propagators in a diagram can be identical, as shown in Figure 3.1.

All the integrals appearing in a single diagram share the same set of propagators $\{D_i\}$, and in a given calculation, it is often the case that many diagrams can be defined with the same set of propagators, or

that diagrams with fewer propagators can be defined such that their set of denominator is a subset of the denominators of some other diagram. This provides us with an organisational principle to sort the integrals appearing in a calculation, which is much needed as the number of scalar integrals in a given calculation often ranges in the thousands. As a result, we define a concept of *Integral Families*, which are also often called *Topologies*, for a reason which we will highlight shortly.

Definition 1. *Let $K = \{k_i\}_{1\leq i\leq L}$ be a finite set of loop momenta, $P = \{p_i\}_{1\leq i\leq E}$ be a finite set of external momenta and $M = \{m_i\}_{1\leq i\leq N}$ be a finite list of masses, with possible redundancies. Let $Q = \{q_i\}_{1\leq i\leq N}$ be a set of momenta taken from $\left(Span_{\mathbb{Q}}(K) \setminus \{0\} + Span_{\mathbb{Q}}(P)\right)$ and $D = \{D_j = (q_j - m_j)\}$ such that*

$$\mathbb{Q}[D,\, P\cdot P,\, M] = \mathbb{Q}[K\cdot K,\, K\cdot P,\, P\cdot P,\, M].$$

*The **Family** or **Topology** defined by D is the $\mathbb{Q}(P\cdot P,\, M)$-vector space generated by Feynman integrals of the form*

$$\int \prod_{l=1}^{L} \frac{d^d k_l}{(2\pi)^d} \frac{\mathcal{N}}{D_1^{n_1}\dots D_N^{n_N}}, \tag{3.5}$$

with $\mathcal{N} \in \mathbb{Q}[K\cdot K,\, K\cdot P,\, P\cdot P,\, M]$ and $n_i \in \mathbb{N}^+$.

The condition on the propagators means that all scalar products in the numerators can be expressed as linear combinations of propagators and external scales. This has a nice consequence as it provides a generating family for our vector space. Indeed, let us consider a family defined by a set of denominators D, then for any numerator $\mathcal{N}$, we can rewrite it as a polynomial of the denominators and the masses:

$$\mathcal{N} = \sum_k q_k D_1^{a_{1k}} \dots D_N^{a_{Nk}}, \tag{3.6}$$

where the q_k are polynomials of the external scales. As a result the integral can be rewritten only in terms of numerators:

$$\int \prod_{l=1}^{L} \frac{d^d k_l}{(2\pi)^d} \frac{\mathcal{N}}{D_1^{n_1}\dots D_N^{n_N}} = \sum_k q_k \int \prod_{l=1}^{L} \frac{d^d k_l}{(2\pi)^d} \frac{1}{D_1^{n_1-a_{1k}}\dots D_N^{n_N-a_{Nk}}}. \tag{3.7}$$

Because the $q_k \in \mathbb{Q}(P\cdot P, M)$, this shows that the integrals of the form

$$I(n_1,\dots,n_N) = \int \prod_{l=1}^{L} \frac{d^d k_l}{(2\pi)^d} \frac{1}{D_1^{n_1}\dots D_N^{n_N}} \tag{3.8}$$

for $\vec{n} \in \mathbb{Z}^N$ form a generating set of the family. We have developed the program `SAIF`, described in Appendix B to express integrals with arbitrary numerators as linear combinations of $I(n_1,\dots,n_N)$ integrals.

Our discussion has shown that it would be sufficient to calculate all the $I(n_1,\dots,n_N)$ to know all the integrals in the family. This, of course does not seem to provide a significant simplification, as there are still infinitely many of them, but we shall see in the next section that we need to calculate much fewer integrals in reality.

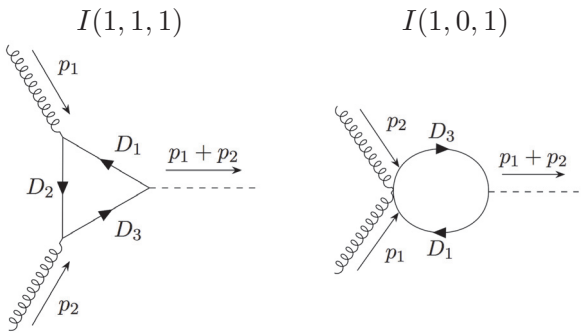


Figure 3.2: Graphical representations of two integrals in the family used for the calculation of gluon fusion at LO

Graphical representation of scalar integrals It is often useful to give a graphical representation for scalar integrals, as, for example, this allows one to easily see when individual integrals from different topologies are actually identical. For integrals of the form $I(n_1, \ldots, n_N)$, we simply represent our integrals as Feynman diagrams that share the same set of propagator denominators. Extra powers of propagators are indicated by dots on the corresponding edges and negative powers or numerators are generically indicated by a dot inside the diagram, or a small note indicating the value of the numerator. Because we will consider QCD processes, we conventionnaly assign gluon lines to massless propagators and either quark or Higgs boson lines to massive propagators. For example, consider the one-loop family of massless bubble integrals defined by $D_1 = k^2 - m^2$, $D_2 = (k+p_1)^2 - m^2$, $D_3 = (k+p_1+p_2)^2 - m^2$ where k is the loop momentum, $p_i^2 = 0$ and $(p_1 + p_2)^2 = m_H^2$. We show in Figure 3.2 how some integrals in the family can be represented. We will use this representation throughout the next chapters to illustrate the loop integrals we considered.

3.1.2 Integration by parts and Master integrals

As we mentionned in the previous subsection, reducing the integrals in a Feynman diagram to the form $I(n_1, \ldots, n_N)$ for a given topology does not reduce the number of integrals we need to evaluate. However, this number can be significantly reduced as we can express *all* the integrals of the form $I(n_1, \ldots, n_N)$ in a given family as linear combination of a finite number of integrals [128]. In other words, families are always finite dimensional vector spaces on the field of rational functions of external scales. While existence theorems can in general be frustratingly abstract, we do know in principle how to reduce our infinite generating set for the family, $\left\{I(n_1, \ldots, n_N), \vec{n} \in \mathbb{Z}^B\right\}$, to a finite generating set, which is conjectured to be free and therefore often called a *basis* of the family, or a set of *master integrals*. Several methods exist to reach a basis, and they are all based on the general idea in linear algebra for reducing a family of vectors to a free family: one finds linear relations between the vectors to eliminate redundant elements. In the case of integral families, finding linear relations is in principle not straightforward, but the *Integration-by-Parts Identities* (IBP) provide a powerful tool that can generate arbitrarily many linear relations between integrals in a given topology [129, 130]. These IBP relations rely on the fact that integrals in dimensional regularization are shift invariant. Indeed, for any regular

enough function $f(k)$, momentum p, and real number h

$$\int d^d k\, f(k) = \int d^d k\, f(k+hp)$$

$$\implies \lim_{h\to 0} \int d^d k\, \frac{f(k) - f(k+hp)}{h} = 0 \tag{3.9}$$

$$\implies \int d^d k'\, \frac{\partial f}{\partial k^\mu}(k') = 0$$

Let us apply this identity on the integrand of one of the integrals in a L-loop topology defined by propagators $D_1, \ldots, D_N$

$$\int \prod_{l=1}^{L} \frac{d^d k_l}{(2\pi)^d} \frac{\partial}{\partial k_i^\mu} \frac{v^\mu}{D_1^{n_1} \ldots D_N^{n_N}} = 0, \tag{3.10}$$

in which we chose to derive with respect to one of the loop momenta k_i and v^μ is either one of the other loop momenta or an external momentum. It is easy to see that evaluating the left-hand side of Equation 3.10 will yield a linear combination of integrals in the family:

$$\int \prod_{l=1}^{L} \frac{d^d k_l}{(2\pi)^d} \frac{\partial}{\partial k_i^\mu} \frac{v^\mu}{D_1^{n_1} \ldots D_N^{n_N}} = \int \prod_{l=1}^{L} \frac{d^d k_l}{(2\pi)^d} \frac{\partial v^\mu}{\partial k_i^\mu} \frac{1}{D_1^{n_1} \ldots D_N^{n_N}} - \sum_{j=1}^{N} \frac{n_j}{D_{n_j}} \frac{\partial D_{n_j}}{\partial k_i^\mu} \frac{v^\mu}{D_1^{n_1} \ldots D_N^{n_N}}. \tag{3.11}$$

The derivative in the first term will yield 1 if $v^\mu = k_i^\mu$ and 0 otherwise and $v^\mu \frac{\partial D_{n_j}}{\partial k_i^\mu}$ is a linear combination of scalar products of loop momenta and external momenta, whence each term takes the form described in Definition 1. We can therefore reduce each term to linear combinations of I integrals and obtain identities of the form

$$\sum_i C_i I(n_{1i}, \ldots, n_{Ni}) = 0. \tag{3.12}$$

Generating these equations for a given family provides the essential tool to obtain a basis and all known techniques are based on IBP relations. The first applications [129, 130] were based on generating IBP relations for general powers of the propagators and solving the system of equations symbolically, thus yielding general expressions for any integral in the family. While this technique is very satisfying from a theoretical point of view, no general algorithm exists that can provide solutions in generic cases. In the following we will present a different method, which has become a staple of the Feynman integral calculation toolbox: the Laporta algorithm [131, 132].

Instead of providing a general solution to the system of equations generated by the IBP relations for general powers of the propagators, this method relies on a total order among $I(n_1, \ldots, n_N)$ integrals, which can be intuitively interpreted as sorting the integrals in terms of complexity. One then builds IBP relations with explicit values for the propagator powers starting from the least complex integrals and moving up the complexity ladder progressively. At each step of the ladder, either all integrals can be expressed in terms of simpler integrals, or some irreducible ones appear that cannot be related to simpler integrals, and these are master integrals. After a number of steps that is dependent on the family, the set of master integral stabilizes and constitutes the set of master integrals. One of the nice

features of this algorithm is that it specifies a uniquely determined basis.

Definition 2. *The Laporta basis for a topology T equipped with a total order $<$ is the set of integrals of the form $I(n_1, \dots, n_N)$ such that no IBP relation can express them as linear combination of smaller integrals.*

While this algorithm can in principle find a complete basis of the family, it is often sufficient to stop before that. In practice, we have a precise list of Feynman integrals we want to simplify, so one can stop when the set of IBP relations generated by the algorithm involves all the integrals we want to simplify.

There is no unique natural way to define a total order in a family. There is however a natural partial organization provided by *sectors*.

Definition 3. *In an integral family defined by the set of denominators $D_1, \dots D_N$, the sector indexed by $\nu \subset [\![1, N]\!]$ is the vector space generated by the $I(n_1, \cdot, n_N)$ whose positive powers are at most ν.*

For example, the integrals of the form $I(0, n_1, n_2)$ all belong to the same sector. Sectors provide a structure on integral that is compatible with many operations:

- If we apply the IBP identity to the integrand in a given sector, the identity only relates integrals within that sector: $\int \frac{d^d k}{(2\pi)^d} \frac{\partial}{\partial k^\mu} \frac{k^\mu}{(k-m^2)((k+p)^2-m^2)}$ does not generate new denominators.
- The same is true for differentiation: $\frac{\partial}{\partial m} I(0, n_1, n_2)$ yields a linear combination of integrals with non-positive first entries

Some sectors are related to each other:

Definition 4. *Given two sectors S and S' indexed by respectively ν and ν', S' is a subsector of S if $\nu \subsetneq \nu'$.*

We can assign a diagrammatic representation to each sector: it is the diagram of the integral where all non-negative powers are set to 0 and all strictly positive powers are set to 1. The diagram of a sector is related to that of its subsectors by pinching the lines with non-positive powers in the subsector.

Integrals with sufficiently many non-positive integer powers are zero, this is in particular the case for integrals with only negative powers, since integrals of polynomials are always scaleless. As a result, it is natural to interpret sectors with more denominators as more complicated.

Definition 5. *Let S and S' be two sectors in a family. If integrals in S have more denominators than the integrals in S', we assign the relation $S' \leq S$.*

Since what we really want to construct is a total order on the integrals $I(n_1, \dots, n_N)$ themselves, we also need to decide on an ordering between sectors. Several choices have been proposed in the literature, and here we present the one described in [133]. Let us consider two integrals $I_1 = I(\vec{n}_1)$ and $I_2 = I(\vec{n}_2)$, we assign $I_1 \leq I_2$ if

1. the sector of I_1 is smaller than the sector of I_2:

$$I(1, 1, 0) < I(1, 1, 1)$$

2. if one sector is not larger, the sum of the positive powers of I_1 is smaller than that of I_2:

$$I(1,1,0) < I(2,0,1)$$

3. if they are equal, the sum of the non-positive powers of I_1 is *larger* than that of I_2:

$$I(1,2,-1) < I(2,-2,1)$$

4. if they are equal, the sequence $(n_{11}\theta(n_{11}),\ldots,n_{1N}\theta(n_{1N}))$ is lexicographically smaller than $(n_{21}\theta(n_{21}),\ldots,n_{2N}\theta(n_{2N}))$, where $\theta(n)$ is 1 if $n>0$ and 0 otherwise:

$$I(1,-2,2) < I(2,-2,1)$$

5. if they are equal, the sequence $(-n_{11}\theta(-n_{11}),\ldots,-n_{1N}\theta(-n_{1N}))$ is lexicographically smaller than $(-n_{21}\theta(-n_{21}),\ldots,-n_{2N}\theta(-n_{2N}))$:

$$I(2,-1,-2) < I(2,-2,-1)$$

.

This provides a total ordering on the integrals, *i.e.* for any I_1, I_2, I_3,

- $I_1 \leq I_2$ or $I_2 \leq I_1$,
- if $I_2 \leq I_1$ and $I_1 \leq I_2$, $I_1 = I_2$,
- if $I_2 \leq I_1$ and $I_3 \leq I_2$, then $I_3 \leq I_1$.

Let us now describe how the algorithm proceeds to build a basis. One starts with a list of integrals $\mathcal{I} = \{I_1,\ldots,I_n\}$. We define a set of replacement rules R initialized at $R=\emptyset$, and an integer K which is initialized at $K=1$.

1. Generate r_K, the set IBP relations with integrands whose powers $n_1,\ldots,n_N$ verify $\sum|n_i| = K$.
2. For each IBP relation in r_k $\sum a_i J_i = 0$, where the J_i are integrals, express them as $J_{\hat{j}} = \sum a_i/a_{\hat{j}} J_i$ where $J_{\hat{j}}$ is the largest integral in the relation.
3. Simplify every possible integral in r_k using R, then $R = R \cup r_k$.
4. If any of the new relations is linearly dependent of the others in R, eliminate it.
5. If every integral in $\mathcal{I}$ figures in the left-hand-side of a relation in R, stop
6. Else $K = K+1$ and repeat.

After this procedure terminates, the set of integrals in the right-hand side of the relations in R constitutes the set of master integrals necessary to express the integrals in $\mathcal{I}$.

Automation is of course crucial for modern real-life applications. For example, we will see in Chapter 5.1 that we need to evaluate more than 1900 integrals in order to express the amplitude for Higgs boson production at NLO in the SMEFT. However, we only need 18 different master integrals in order to express them. Not only is it unimaginable to evaluate these 1900 integrals individually, but using IBP reductions to reduce them to master integrals by hand is also unrealistic.

Luckily, the Laporta algorithm has been implemented in a number of public codes, like `FIRE` [134, 135], `AIR` [136], `LiteRed` [137], or `Kira` [138] which allow one to automatically reduce integrals in a topology to a set of master integrals with very little user input. In practice, in the projects we will present in the following chapters, we used both `LiteRed` and `FIRE` to reduce integrals to master integrals.

3.2 Evaluating scalar integrals

In this section, we will describe the different methods we have used to evaluate Feynman integrals. Calculating multiloop integrals is more of an art than a science and no systematic, algorithmic method is known to obtain them. Among the modern methods, that of differential equations is probably the most popular, due to its many successes in deriving the most complicated Feynman integrals, such as those relevant for vector boson pair production at NNLO [139]. This method has the particularity of relying not on the direct calculation of the integrals, but on the derivation of a system of differential equations verified by the integrals in a basis, which can then be solved as an expansion ϵ [140–142].

In the projects presented in this thesis, we have however used more traditional methods to evaluate integrals, which rely on direct integration in the Feynman parameter representation. We will first discuss the general framework of Feynman parameters and show how one can industrialize the approach presented in quantum field theory textbooks. We will then discuss the method of regions, which is a general scheme to obtain approximations of integrals in certain limits. Finally we will discuss the numerical integration of Feynman integrals based on the method of sector decomposition, which we used in both the projects presented in this thesis to cross check our analytic results.

3.2.1 Feynman parameters

The classical approach to Feynman integrals presented in many QFT textbooks is based on Feynman parameters, which is a technique to transform loop integrals expressed in terms of loop momenta into integrals over rational functions, which are usually easier to deal with. This approach is based on two identities:

- The value of tadpole integrals:

$$\int \frac{d^d k}{(2\pi)^d} \frac{1}{(\mathcal{U}k^2 - \mathcal{F})^n} = (-1)^n i (4\pi)^{-d/2} \frac{\Gamma(n - d/2)}{\Gamma(n)} \frac{\mathcal{U}^{-d/2}}{\mathcal{F}^{n-d/2}} \tag{3.13}$$

- The *Feyman Trick*:

$$\frac{1}{D_1^{n_1}\dots D_N^{n_N}} = \frac{\Gamma(n)}{\prod_i \Gamma(n_i)} \int_0^1 \prod_i x_i^{n_i-1} dx_i \frac{\delta\left(1-\sum x_i\right)}{\left(\sum x_i D_i\right)^n}, \tag{3.14}$$

  where $n = \sum n_i$. The parameters x_i are usually referred to as *Feynman parameters*.

One can use the Feynman trick to convert a L-loop integral into a tadpole integral:

$$I(n_1,\dots,n_N) = \frac{\Gamma(n)}{\prod \Gamma(n_j)} \int \prod_i \frac{d^d k_i}{(2\pi)^d} \int \prod_j dx_j x_j^{n_j-1} \delta(1-\sum n_k) \frac{1}{(\sum x_k D_k)^n}. \tag{3.15}$$

Since the D_i can be expressed as $(a_{il}k_l + b_{iq}p_q)^2 - m_i^2$, where the p_q are the external loop momenta and the m_i are masses, the denominator of the Feynman-parametrized integral can be written as

$$\sum_i x_i D_i = A^{lm} k_m k_l + 2P^l k_l + M^2 = K^T \cdot A \cdot K + 2K \cdot P + M^2, \tag{3.16}$$

where A is a symmetric matrix with scalar entries, K is the L-dimensional vector of loop momenta, and P and M are respectively a vector and a scalar independent of the loop momenta. We can use the shift invariance of loop integrals to eliminate the linear term by applying the change $K = K' - A^{-1}\cdot P$:

$$\sum_i x_i D_i = K'^T A K' + \tilde{M}^2. \tag{3.17}$$

We can then diagonalize A by performing an orthogonal transformation $K' = RK''$ with $K'' = (\tilde{k}_1,\dots,\tilde{k}_N)$, and define $\tilde{A} = R^T A R = \text{Diag}(\tilde{A}_1,\dots\tilde{A}_N)$. The integral over loop momenta we need to compute takes the form

$$\tilde{I} = \int \prod_i \frac{d^d \tilde{k}_i}{(2\pi)^d} \frac{1}{(\tilde{A}_1\tilde{k}_1^2 + \dots + \tilde{A}_N \tilde{k}_N^2 - \tilde{M}^2)^n}, \tag{3.18}$$

Which is something we can do iteratively. Let us define $\mathcal{U}_i = \tilde{A}_i$ and $\mathcal{F}_i = -\sum_{j>i} A_j \tilde{k}_j^2 + \tilde{M}^2$. After performing l integrals, we find

$$\tilde{I} = (-1)^{\sum_{i\le L} n_i} (i(4\pi)^{-d/2})^l \frac{\Gamma(n-ld/2)}{\Gamma(n)} \int \prod_{i>l} \frac{d^d\tilde{k}_i}{(2\pi)^d} \frac{(\mathcal{U}_1\dots\mathcal{U}_l)^{-d/2}}{(\mathcal{U}_{l+1}\tilde{k}_{l+1}^2 - \mathcal{F}_l)^{n-ld/2}}, \tag{3.19}$$

which terminates at

$$\tilde{I} = (-1)^n (i(4\pi)^{-d/2})^L \frac{\Gamma(n-Ld/2)}{\Gamma(n)} \frac{(\mathcal{U}_1\dots\mathcal{U}_L)^{-d/2}}{(\mathcal{F}_l)^{n-Ld/2}}. \tag{3.20}$$

Note that, in the numerator of $\tilde{I}$, $\prod \mathcal{U}_i$ is the determinant of the matrix A and $\mathcal{F}_l = \tilde{M}^2 = M^2 - P^T A^{-1} P$. It is useful to re-shuffle the ratio $\frac{\mathcal{U}_1\dots\mathcal{U}_L}{(\mathcal{F}_l)^{n-Ld/2}}$ into $\frac{\mathcal{U}^{n-(L+1)d/2}}{\mathcal{F}^{n-Ld/2}}$, where $\mathcal{U} = \det(A)$ and

$\mathcal{F} = \det(A)\mathcal{F}_l$. Using this notation, we can give a general form for the L-loop integral $I(n_1, \dots, n_N)$:

$$I(n_1, \dots, n_N) = (-1)^n (i(4\pi)^{-d/2})^L \frac{\Gamma(n - Ld/2)}{\prod \Gamma(n_i)} \int_0^\infty \left(\prod x_i^{n_i - 1} dx_i\right) \delta(1 - \sum x_i) \frac{\mathcal{U}^{n-(L+1)d/2}}{\mathcal{F}^{n-Ld/2}}, \tag{3.21}$$

where the integration can be taken from 0 to ∞ because the delta function forces $x_i \leq 1$, and $\mathcal{U}$ and $\mathcal{F}$ can be determined from the denominator of the integral directly after the insertion of the Feynman parameters. Due to their definition that can be determined directly from the expression of the denominator of the integral after inserting the Feynman parameters, expressing any integral $I(n_1, \dots, n_N)$ in the form of Equation 3.21 is straightforward to do algorithmically and automated tools to do so from the definition of a topology, such as `UF` [143].

The choice of $\mathcal{U}$ and $\mathcal{F}$ to parametrize Feynman integrals has several advantages. First of all, both of them are polynomials in the Feynman parameters with uniform degree. Indeed, the matrix A has a polynomial dependence on the x_i with rational coefficients, making $\mathcal{U} = \det(A)$ a polynomial of uniform degree L, while $\mathcal{F}$ has the following expression:

$$\begin{aligned} \mathcal{F} &= \det(A)\mathcal{F}_l = \det(A)\left(M^2 - P^T A^{-1} P\right) \\ &= \det(A) M^2 - P^T \mathrm{Adj}(A) P, \end{aligned} \tag{3.22}$$

where $\mathrm{Adj}(A)$ is the adjugate matrix of A, constructed out of minors of A, and is expressed as a uniform polynomial of degree $L - 1$ in the entries of A. Because $P \cdot K$ is of uniform degree 1 in the Feynman polynomials and K is of degree 0, P is of uniform degree 1, the second term in $\mathcal{F}$ has a uniform degree $L + 1$, while M^2 is of degree 1, so that $\mathcal{F}$ overall of uniform degree $L + 1$.

Another advantage of using these polynomials is that they can easily be constructed by hand from the topology [144]. Indeed, let $\mathcal{G}$ be the graph of the integral $I(1, \dots, 1)$ in the sense defined in Section 3.1.1 seen as a set of edges e_i labelled by the Feynman parameters mutliplying the associated propagator $\{(e_i, x_i)\}$. The polynomial $\mathcal{U}$ can be constructed by considering $\mathcal{T}_1$, the spanning 1-forest of $\mathcal{G}$, which is to say the set of connected tree diagrams that can be constructed by removing L edges from $\mathcal{G}$:

$$\mathcal{U} = \sum_{T \in \mathcal{T}_1} \prod_{i | e_i \notin T} x_i, \tag{3.23}$$

that is $\mathcal{U}$ is the sum of monomial that are the products of L Feynman parameters associated to propagators which can be removed to make $\mathcal{G}$ into a tree. We can similarly construct $\mathcal{F}$ out of $\mathcal{T}_2$, the set of 2-trees of $\mathcal{G}$, which is to say the set of disconnected pairs of trees that can be constructed by removing $L + 1$ edges from $\mathcal{G}$. Each such pair of trees T has an associated invariant $q^2(T)$, where $q(T)$ is the momentum flowing from one tree to the other. The polynomial $\mathcal{F}$ is then constructed in the following way:

$$\mathcal{F} = -\sum_{T \in \mathcal{T}_2} q^2(T) \prod_{i | e_i \notin T} x_i + \mathcal{U} \sum_i m_i^2 x_i, \tag{3.24}$$

where m_i is the mass in propagator D_i.

Cheng-Wu Theorem The Cheng-Wu theorem is the following [145]

Theorem 1. *Let $F : \mathbb{R}^N \to \mathbb{R}$ be a homogeneous function of degree $-N$, then for any non-empty $\nu \subset [\![1, N]\!]$,*

$$\begin{aligned}&\int_0^\infty \left(\prod_{i=0}^N dx_i\right) \delta\left(1 - \sum_{i=0}^N x_i\right) F(x)\\ &= \int_0^\infty \left(\prod_{i=0}^N dx_i\right) \delta\left(1 - \sum_{i\in\nu} x_i\right) F(x).\end{aligned} \tag{3.25}$$

While it seems unsignificant, this theorem is a great help in practical calculation of loop integrals through direct integration, as choosing a smart combination of variables that will be fixed by the delta function can substantially simplify integrals. It is nearly always most useful to have only one variable leftover in the delta function, so that it is fixed to one and one is left with an integral over a rational function between 0 and ∞. Proving the Cheng-Wu theorem can be done shortly by using a "nontrivial value of 1":

$$1 = \int_0^\infty \frac{d\eta}{\eta} \delta\left(1 - \frac{1}{\eta}\sum_\nu x_i\right). \tag{3.26}$$

Let us insert this expression in our integrand

$$\int_0^\infty \left(\prod_{i=0}^N dx_i\right) \delta\left(1 - \sum_{i=0}^N x_i\right) F(x) \tag{3.27}$$

$$= \int_0^\infty \left(\prod_{i=0}^N dx_i\right) \int_0^\infty \frac{d\eta}{\eta} \delta\left(1 - \frac{1}{\eta}\sum_\nu x_i\right) \delta\left(1 - \sum_{i=0}^N x_i\right) F(x), \tag{3.28}$$

and apply the change of variable $x_i = \eta\alpha_i$:

$$\int_0^\infty \left(\prod_{i=0}^N dx_i\right) \int_0^\infty \frac{d\eta}{\eta} \delta\left(1 - \frac{1}{\eta}\sum_\nu x_i\right) \delta\left(1 - \sum_{i=0}^N x_i\right) F(x) \tag{3.29}$$

$$= \int_0^\infty \left(\prod_{i=0}^N d\alpha_i\right) \int_0^\infty \frac{d\eta}{\eta} \delta\left(1 - \sum_\nu \alpha_i\right) \delta\left(1 - \eta\sum_{i=0}^N \alpha_i\right) F(\alpha) \tag{3.30}$$

$$= \int_0^\infty \left(\prod_{i=0}^N d\alpha_i\right) \delta\left(1 - \sum_\nu \alpha_i\right) F(\alpha), \tag{3.31}$$

$$\tag{3.32}$$

thanks to the fact that

$$\int_0^\infty \frac{d\eta}{\eta} \delta\left(1 - a\eta\right) = \int_0^\infty \frac{d\eta}{a\eta} \delta\left(1/a - \eta\right) = 1. \tag{3.33}$$

The two key elements for this proof to work are the homogeneity of the loop integrand, which is always verified for Feynman integrals and the fact that the integration is performed from 0, to ∞, to ensure that all multiplicative changes of variables leave the bounds unchanged.

3.2.2 Approximating integrals: the method of regions

Evaluating Feynman integrals through direct integration of the Feynman parameter representation or by any other method is often a difficult task, especially for those integrals that depend on several parameters. It is however often the case that the general dependence on such external scales is more than we need to study the phenomenology of interesting physical processes and it can be a successful approach to obtain approximate results for observables in specific limits which simplify the calculation. In quantum field theory, this idea essentially has two realizations which are tightly related. The first approach is based on effective field theories, in which one derives a Lagrangian that approximates reality in a specific regime which automatically generates observables as a power series in some small parameters. The second approach to approximating observables in given limits is to perform the expansion at the amplitude level. This type of calculation is actually necessary for EFT that are used as approximations of a known UV-complete theory[1], such as the two we described above. As approximations of a known theory, the value of their paramaters can be derived from the parameters of the full theory. This is done by the process of *matching*, which we discussed in the context of the SMEFT in Section 2.4. To obtain the value of the EFT coefficients, one calculates the prediction for a given observable in the EFT, which is automatically expressed as a power series in the expansion parameter, and in the full theory, which one then needs to expand. In simple cases, it is possible to calculate such observables in the full theory exactly. However, it can also be the case that it is a better choice to expand the full theory prediction at the integrand level. This is what we do in Chapter 4, in which we extract a HEFT coefficient by performing the expansion by region of a SM two-loop amplitude.

The method of choice to perform such an expansion is the *method of regions* [146]. It provides a systematic way to do expansions of loop integrals at the integrand level, thus simplifying the integrals one has to calculate. It can be used, as we will see in Chapter 4, as a way to relate a UV-complete theory to an EFT, or as a way to bypass the use of an EFT. It is indeed not always necessary to develop a full EFT, which can get particularly involved, if one is only interested in a limit of one single observable. The counterpart, however, is that every new observable requires to be expanded from scratch, while performing the matching once on simple observables and then doing all calculations in an EFT can be a big advantage if one needs to make many predictions in a given limit.

The method of regions is motivated by the observation that making expansions at the loop integrand level can be a tricky business. Indeed, let us consider an off-shell bubble integral with massive propagators already discussed in [146]:

$$I(p,m) = \int d^dk \frac{1}{((k+p/2)^2-m^2)((k-p/2)^2-m^2)}, \tag{3.34}$$

where, m is a mass and $p^2 \neq 0$. If we want to consider a threshold limit, in which $p^2 \to 4m^2$, we can define a parameter $y = m^2 - p^2/4$ and try to obtain the integral as an expansion in the limit $y/p^2 \to 0$. The naive approach would be to expand the integrands in powers of y to obtain in power series:

$$I(p,m) \underset{\text{naive}}{=} \int d^dk \frac{1}{(k^2-k\cdot p)(k^2+k\cdot p)} + y\int d^dk \frac{2k^2}{(k^2-k\cdot p)^2(k^2+k\cdot p)^2} + \mathcal{O}(y^2)\,. \tag{3.35}$$

[1] Often called *top-down EFT*

However, if we look at the exact formula for $I(p,m)$ and expand, we realize that we are missing terms that we can never obtain with the power expansion of the integrand:

$$\begin{aligned} I(p,m) &= i(\pi)^{d/2}\Gamma(\epsilon)y^{-\epsilon}\,{}_2F_1\left(\frac{1}{2},\epsilon,\frac{3}{2},-\frac{q^2}{4y}\right) \\ &= \frac{1}{\epsilon}+2\frac{4y}{p^2} \\ &+\log\left(\frac{4y}{p^2}\right)-\frac{2\pi\sqrt{y}}{p}+\dots \end{aligned} \tag{3.36}$$

The naive expansion correctly reproduces the first line of the expansion of the full result, but will never be able to reproduce the logarithm and the term proportional to $\sqrt{y}$. This problem can be explained by the fact that our naive expansion blindly considered that the expansion parameter is the only small parameter, while the loop momentum is integrated from $-\infty$ to $+\infty$, therefore taking both small and large values.

The method of region is a formalization of the natural idea following this observation, that one should divide the integration range into different parts where we know how to treat the loop momentum in our integrand expansion. The original formulation of the idea [146] was based on the momentum space formulation of loop integrals and consisted of performing the following steps:

1. Separate your integral as a sum of pieces where the components of the loop momenta scale like small or large parameters.

2. Expand each piece in powers of the small quantities.

3. Integrate each piece over the complete range (from $-\infty$ to $+\infty$).

The crucial element in this sequence is step 3 which seems problematic: if we consider the loop momentum to scale in a definite fashion and we then integrate over the whole range, we restore the fact that our loop momentum goes from small to large, and indeed if we exchange steps 2 and 3, we get $n\times\sum_{\text{naive}} x^i I_i$, where n is the number of regions and the expansion $\sum_{\text{naive}} x^i I_i$ in the small parameter x is exactly the naive integrand expansion which we argued can miss contributions. Indeed, steps 2 and 3 do not commute, which can be understood from the fact that scaleless dimensionally-regularized integrals are zero. When we expand the integrand, we are factorizing some scale or combination of scales out of the integral. As a result, in each region, the integration picks up only these scales that remain in the propagators even if we integrate over the whole range of loop momenta. Let us look at how this appears in our bubble integral example. We have already discussed the contribution of the region where we consider $k^2\gg y$ where the process we discussed was clear: at each order in the expansion, the small quantity y factorizes out of the integrand, which can therefore only pick a dependence on p^2. The only other region that contributes is the *potential region* which underlies the non-relativistic nature of our expansion, since in pair production, $p^2\simeq 4m^2$ corresponds to producing particles near rest. Let us pick a frame in which $p=(p_0,0,0,0)$ and specify the scaling $k\sim(y/p_0,\sqrt{y},\sqrt{y},\sqrt{y})$. In the integrand of

$I(p,m)$, we can neglect the terms proportional to k_0^2 and get the following integral:

$$\int dk_0 d^{d-1}\vec{k} \frac{1}{(\vec{k}^2 + p_0 k_0 + y)(\vec{k}^2 - p_0 k_0 + y)}. \tag{3.37}$$

We can use the Cauchy theorem to integrate over k_0 and get

$$\frac{i\pi}{\sqrt{p^2}} \int d^{d-1}\vec{k} \frac{1}{\vec{k} + y}, \tag{3.38}$$

which yields the missing piece from the expansion of the exact result. In this integral, we integrated over k_0 by extracting the residue of the poles of the propagator at $k_0 = \pm\dfrac{\vec{k}^2 + y}{q_0}$, showing that, despite integrating over the whole range, we only collect contributions from the region where k_0 is soft. Similarly, the integral over $\vec{k}$ only depends on y, so that despite integrating all the way to infinity, it is the scale y that controls the integral.

Of course, while we recover the correct expansion, this discussion is very hand-wavy, which betrays the fact that the method of regions is not proved to be correct. It has however become a major technique in the toolbox for the evaluation of Feynman integrals and has successfully provided many results, which have always matched the explicit expansion of exact results when available. In modern multi-loop calculations, however, the formulation described above in terms of loop momenta has been mostly abandonned in favor of Feynman parameters, which is much more systematic and avoids several issues of the loop-momentum formulation, such as the possibility of double counting and of obfuscation of regions due to the shift invariance of the loop momenta [147][2].

In this formulation a region is defined by a definite scaling of Feynman parameters in the small parameter y:

$$x_i \sim y^{\kappa_i} \tag{3.39}$$

where κ_i is some integer. In each region we compute the expansion of the integrands and then integrate the Feynman parameters between 0 and ∞. Mechanically, this can be done by introducing a auxiliary parameter ρ to keep track of the power dependence of each term in the integrand: we assign a scaling in ρ to each scale, a scaling in ρ to each Feynman parameter and then express the integrand as a power series in ρ, which we can set to one after the expansion. This trick is both a useful bookkeeping device to track the scaling of each monomial in $\mathcal{U}$ and $\mathcal{F}$, and a nice technical tool to define the expansion algorithmically.

While the expansion of the polynomials in the integrand is straightforward, the treatment of the delta function in the Feynman parametrization is less obvious. We can actually keep it as $\delta\left(\sum_{i\in\nu} x_i - 1\right)$, for any subset ν of the indices even after the rescaling and expansion. There are two ways of showing that this is true.

[2]The shift invariance means that it can be that the property of being soft or hard must be associated to linear combinations instead of individual momenta, which might not have a well-defined scaling. As a result, if the wrong parametrization is chosen, it can be difficult to find regions.

- One can start from the Schwinger parametrization of integrals, which is also formulated in terms of the Symanzik polynomials but does not contain delta functions. In this parametrization, one can expand the integrand in powers of ρ and then apply the classical trick [133] to switch to Feynman parametrization. As this trick only relies on the homogeneity of the Symanzik polynomial, which is conserved in the expansion, one reaches a Feynman-parametrized integral with expanded polynomials and the usual delta function, independent of ρ.
- If we start from the Feynman parametrization, we can use the Cheng-Wu theorem to keep only those x_i in the integral that will be rescaled by ρ^0. After the expansion, which leaves the delta function unchanged, we have new integrals with new homogeneous integrands, which is sufficient to use the Cheng-Wu theorem and restore any rescaled x_i into the delta function.

We are now equipped to write the expansion of a generic integral

$$I = I_0 \int (dx_i\, x_i^{a_i})\, \delta(1 - \sum x_i) \frac{\mathcal{U}^p}{\mathcal{F}^q}. \tag{3.40}$$

After rescaling by ρ, the polynomials are ordered in the following way:

$$\mathcal{U} = \rho^{\kappa_{\mathcal{U}}} \sum_{i=0}^{n_U} \rho^i U_i \tag{3.41}$$

$$\mathcal{F} = \rho^{\kappa_{\mathcal{F}}} \sum_{i=0}^{n_F} \rho^i F_i, \tag{3.42}$$

and we can write the expansion of a generic Feynman integral I. After setting $\rho = 1$, the expansion takes the form

$$I = I_0 \int (dx_i\, x_i^{a_i})\, \delta(1 - \sum x_i) \left(\frac{U_0^p}{F_0^q} + \frac{U_0^{p-1}}{F_0^{q-1}} (pU_1F_0 - qU_0F_1) + \dots \right), \tag{3.43}$$

where the neglected terms are all of the form $P\dfrac{U_0^a}{F_0^b}$ where P is a polynomial.

Of course, a key point in working in this representation is to find the relevant regions as, even if one restricts oneself to $\kappa_i \in \{0, 1\}$, there are 2^N possible sectors, where N is the number of propagators, while in reality most are zero. Thankfully, we know an *a priori* criterium that can be applied algorithmically to find relevant regions [148] using geometry.

A sufficient condition for a Feynman integral to be zero is that it be scaleless. If we consider a Feynman parametrized integral I and apply the change of variable $x_i \to \alpha^{\kappa_i} x_i$ with at least one $\kappa_i = 0$, and find that this transforms the integral homogeneously, $I \to \alpha^{\kappa_I} I$, the integral is obviously trivial since it is independent of α. For this behavior to happen, it is necessary that, under such a change, both Symanzik polynomials transform homogeneously:

$$\begin{aligned} \mathcal{U} &\to \alpha^{\kappa_U} \mathcal{U} \\ \mathcal{F} &\to \alpha^{\kappa_F} \mathcal{F}, \end{aligned} \tag{3.44}$$

which, because $\mathcal{U}$ and $\mathcal{F}$ are polynomials, is equivalent to $\mathcal{U}\mathcal{F}$ transforming homogeneoulsy.

Furthermore, if an integral is scaleless, multiplying its integrand by a polynomial will also yield a scaleless integral since each monomial is homogeneous. As a result, remembering the structure of the integrand in the expansion by region described in Equation 3.43, for a region to be irrelevant, it is sufficient to show that $U_0 F_0$ is homogeneous under rescaling of a subset of the parameters. We will now show that we can give a geometric characterization of the regions that generate such homogeneous polynomials.

Let us switch to a geometric description of this polynomial: to $\mathcal{UF}$, we substitute a set points $P[\mathcal{UF}]$ in $N+1$-dimensional space associated to its monomials:

$$s_i x_1^{v_1} \dots x_N^{v_N} \to \vec{v} = (v_0, v_1, \dots, v_N)^T , \tag{3.45}$$

where the s_i is an external scale that scales like ρ^{v_0}. In this picture, in a region r defined by $x_j \to \rho^{r_j} x_j$, each monomial is transformed in the following way:

$$\begin{aligned} s_i x_1^{v_1} \dots x_N^{v_N} &\to \rho^{v_0 + r_1 v_1 + \dots r_N v_N} s_i x_1^{v_1} \dots x_N^{v_N} && (3.46) \\ &= \rho^{\vec{r}\cdot\vec{v}} s_i x_1^{v_1} \dots x_N^{v_N} , && (3.47) \end{aligned}$$

where $\vec{r} = (1, r_1, \dots, r_N)$. Let us consider the convex hull of $P[\mathcal{UF}]$, $H[\mathcal{UF}]$, that is the convex polytope whose vertices are the points of $P[\mathcal{UF}]$. Because $\mathcal{UF}$ is of dimension $2L+1$, where L is the loop order, any vector $\vec{v}$ in $P[\mathcal{UF}]$ verifies $v_1 + \dots + v_N = 2L+1$, meaning that $H[\mathcal{UF}]$ has dimension N. More precisely, $H[\mathcal{UF}]$ is in the hyperplane orthogonal to $(0, 1, \dots, 1)^T$. We are always free to translate a region vector $\vec{r}$ in the direction $(0, 1, \dots, 1)^T$ since this corresponds to uniform rescalings, so we can chose $\vec{r} \perp (0, 1, \dots, 1)^T$, which is an irrelevant direction when discussing the scaling behavior of $\mathcal{UF}$.

If we consider a region defined by $\vec{r}$, the leading-order polynomial in the expansion of $\mathcal{UF}$ corresponds to a set of points $\{\vec{v}_i\}_{i\in\nu}$ in $P[\mathcal{UF}]$ whose monomials have the same scaling, which implies that they belong to a hyperplane V orthogonal to $\vec{r}$: $(\vec{v}_i - \vec{v}_j)\cdot\vec{r} = 0$. All the neglected terms will correspond to points $\vec{w}$ such that $\vec{w}\cdot\vec{r} > \vec{v}_i\cdot\vec{r}$, and we can decompose $\vec{w} = \vec{v}_i + a\vec{r} + \vec{w}^{\perp\vec{r}}$ such that $\vec{r} \perp \vec{w}^{\perp\vec{r}}$ and $a > 0$. As a result, all the points correponding to neglected terms are *above* V in the $\vec{r}$ direction which means that the intersection of V and $H[\mathcal{UF}]$ is the face[3] of $H[\mathcal{UF}]$ Φ orthogonal to $\vec{r}$ such that $\vec{r}$ points *into* $H[\mathcal{UF}]$.

The basis of the method used by `ASY` [149] is that if Φ is not a facet, that is a maximal-dimensional face of $H[\mathcal{UF}]$, then the associated region yields a scaleless integral, which leaves a finite, and usually small, number of regions to test, defined by the orthogonal vectors of the facets of $H[\mathcal{UF}]$. Let us prove this claim by considering a region $\vec{r}$ orthogonal to a face ϕ which is not a facet. We know that $H[\mathcal{UF}]$ is of dimension N, so ϕ has dimension at most $N-2$. In our $N+1$-dimensional space, we know two explicit vectors orthogonal to ϕ: $\vec{r}$ and $(0, 1, \dots, 1)^T$, and we can find a third one, $\vec{w}$, which we can choose to be of the form $(0, w_1, \dots, w_N)^T$ with at least one $w_i = 0$. This means that there is a nontrivial rescaling $x_i \to \alpha^{w_i} x_i$ that transforms both $\mathcal{U}$ and $\mathcal{F}$ homogeneously, making the integral scaleless.

[3] In N-dimensions, a face is any polytope in the surface of $H[\mathcal{UF}]$. Faces with maximal dimensions are called *facets*

Diagrammatic interpretation of the method of regions Because each Feynman parameter is associated to one propagator in the loop-momentum representation of the integral, there is a nice diagrammatic interpretation of the expansion by region which makes the connection to EFT calculations and to the technique of expansion by subgraphs, which is an rigorous expansion method, but is very limited as it only works in Euclidean space. The expansion by region is therefore not equivalent to the expansion by subgraph, but their language is close enough to suggest an intuitive graphical interpretation of the method of regions.

Assigning a scaling to the Feynman parameters in the method of regions is the analog of considering the associated propagator as either being small or large, which is what is done in the expansion by subgraphs.

In this technique, one defines a set of hard and soft external momenta (internal masses are supposed to be small) and the expansion of an integral I whose graphical representation is Γ is organized in the following way:

$$I = \sum_{\gamma} I_{\Gamma/\gamma} \circ \mathbf{T} I_{\gamma}, \tag{3.48}$$

where the sum runs over asymptotically irreducible subgraphs of Γ (1-particle irreducible subgraph that contains all the vertices connected to external particles with hard momenta). $I_{\Gamma/\gamma}$ is the integral associated to the graph in which γ has been reduced to a point. I_{γ} is the integrand associated to γ, $\mathbf{T}$ denotes the operator associating expansion in powers of the small momenta and masses – which yields integrands whose non-trivial dependence is limited to the hard momenta – and the $\circ$ symbol indicates that we insert this expansion under the integral sign of $I_{\Gamma/\gamma}$.

The interpretation of this expansion is quite straightforward: by assigning each propagator to either γ or Γ/γ, we define it to be either *hard* or *soft*, which is reflected by the fact that the soft scales are factored out of the integration in γ and that Γ/γ is independent of the hard scales due to the conditions on asymptotic subgraphs. Because this interpretation is based on graphs, it is also clear how one should diagrammatically represent this expansion: the integral is a sum of diagrams in which each hard subgraph γ is reduced to a single point. The integral over the expanded propagators in γ is completely independent of the soft scales, and this can be interpreted as the fact that they describe very hard, short-distance interactions, which can be described as local interactions from the point of view of the soft physics.

A similar interpretation can be made in the Feynman parametric expansion: in a given region, Feynman parameters which are assigned a trivial scaling in ρ can be considered as hard while those that contain positive scaling can be considered as soft, and we can represent the associated integral with the corresponding hard propagators contracted to a point. The integrals does not obviously take a structure reminiscent of Equation 3.48, but experience shows that the integral over hard parameters is simplified.

It is hard not to want to make a connection between this diagrammatic description and the EFT expansion of amplitudes: heavy particles are integrated out of the theory and loops of such particles are described by new local interactions whose coefficients are computed by evaluating these loops in the infinite mass limit. The correspondance is of course not directly one-to-one since the EFT description works at the amplitude level while we are discussing individual integrals inside the amplitude, but we shall see in Chapter 4 that it can be a useful guide to understanding the matching between the SM

and the HEFT. In particular, we will observe the factorization of an integral appearing in the SM into a tadpole multiplied by an integral appearing the HEFT, corresponding to the picture described by Equation 3.48.

3.2.3 Numerical evaluation of integrals

In this final section exposing the modern techniques for evaluating Feynman integral, we discuss briefly the method of numerical evaluation by sector decomposition. The goal is not to provide a fully comprehensive description of the method, but to rapidly describe the basics of this technique, which was used during this thesis as a cross-checking tool for our analytic results.

A systematic approach to numerically evaluating Feynman integrals using Monte Carlo methods was proposed in [150–152]. A number of programs have been developped that can evaluate Feynman integrals in this fashion [153–157], in particular the public tools `SecDec` [156] and `FIESTA` [157], which can evaluate integrals for both Euclidean and physical kinematics. They are often used as a check of analytic calculations, but have also produced major results where integrals were not known analytically, as was the case in the case of the evaluation of the NLO corrections to Higgs boson pair production by the `SecDec` collaboration [158], which has been an impressive *tour de force* of the recent years. In the projects described in this thesis, we used both `FIESTA` and `SecDec` to check our analytic results, and they have been instrumental in giving us confidence in the validity of our results.

A major issue with the numerical integration of integrals is the presence of divergences. In analytic calculations, we regulate such divergences using dimensional regulation in which we set the number of dimensions $d = 4 - 2\epsilon$ and then express our integrals as a Laurent series in ϵ. Let us for example consider a massless bubble integral $B(1,2)$ with one propagator squared

$$\begin{aligned} B(1,2) &= \int \frac{d^dk}{i\pi^{d/2}} \frac{1}{k^2((k+p)^2)^2} = \left(p^2\right)^{-\epsilon-1} \frac{\Gamma(1-\epsilon)\Gamma(-\epsilon)\Gamma(\epsilon+1)}{\Gamma(1-2\epsilon)} \\ &= -\frac{1}{p^2\epsilon} + \frac{\log\left(p^2\right)+\gamma}{p^2} + \mathcal{O}(\epsilon)\,, \end{aligned} \tag{3.49}$$

which we can represent as the Feynman parameter integral

$$B(1,2) = \int_0^\infty dx_1 \left(x_1+1\right)^{2\epsilon-1} \Gamma(\epsilon+1) \left(p^2 x_1\right)^{-1-\epsilon} = \Gamma(1+\epsilon) \left(p^2\right)^{-1-\epsilon} \int_0^\infty dx_1 b(x_1,\epsilon). \tag{3.50}$$

Because we need to express our result as a Laurent series in ϵ, we need to keep the dimensional regulator explicit and evaluate the coefficients of the series numerically. Of course, naively expanding $b(x_1,\epsilon)$ in powers of ϵ before integration breaks the regulation:

$$b(x_1,\epsilon) = \frac{1}{x_1\left(1+x_1\right)} + \mathcal{O}(\epsilon)\,, \tag{3.51}$$

which is the original unregulated integral, since regularization and integration do not commute. We can however treat the $x_1^{-1-\epsilon}$ as a distribution and expand it in the following way:

$$x_1^{-1-\epsilon} = -\frac{1}{\epsilon}\delta(x) + \sum_n \frac{\epsilon^n}{n!} \left(\frac{\log^n(x)}{x}\right)_+ \tag{3.52}$$

where we use the plus distribution

$$\int dx f(x) \left(\frac{\log^n(x)}{x}\right)_+ = \int (f(x) - f(0)) \left(\frac{\log^n(x)}{x}\right). \tag{3.53}$$

Inserting this expansion in the integrand yields a series of convergent integrals that we can then further expand in powers of ϵ and evaluate numerically. However, this is still not sufficient to evaluate generic integrals when the divergences appear in the limit where several parameters go to 0. Let us for example consider the integral

$$\int_0^1 dx_1 \int_0^1 dx_2 \frac{1}{(x_1+x_2)^{2+\epsilon}}. \tag{3.54}$$

The fact that a logarithmic divergence appears in the corner of the two-dimensional integration space $(x1, x2) \to (0, 0)$ makes the expansion of the denominator insufficient to have a regulated integral

$$\int_0^1 dx_1 \int_0^1 dx_2 \frac{1}{(x_1+x_2)^{2+\epsilon}} \to \frac{1}{\epsilon} \int_0^1 dx_2 \int_0^1 dx_1 \frac{\delta(x_1+x_2)}{(x_1+x_2)} + \dots \tag{3.55}$$

This motivates the need for *sector decomposition* which is a technique that extracts divergences from the corners of multi-dimensional intregration spaces and factorize them as one-dimensional divergences that can be regulated using Equation 3.52. Let us first see how it works on our example.

We can cut the integration region in the following way:

$$\int_0^1 dx_1 \int_0^1 dx_2 \frac{1}{(x_1+x_2)^{2+\epsilon}} = \int_0^1 dx_1 \int_0^{x_1} dx_2 \frac{1}{(x_1+x_2)^{2+\epsilon}} + \int_0^1 dx_2 \int_0^{x_2} dx_1 \frac{1}{(x_1+x_2)^{2+\epsilon}}, \tag{3.56}$$

and in each, factorize the divergence by remapping the integration to the unit square

$$\int_0^1 dx_1 \int_0^1 dx_2 \frac{1}{(x_1+x_2)^{2+\epsilon}} = \int_0^1 dx_1 \int_0^1 d\tilde{x}_2 \frac{1}{x_1^{1+\epsilon}(1+\tilde{x}_2)^{2+\epsilon}} + \int_0^1 dx_2 \int_0^1 d\tilde{x}_1 \frac{1}{x_2^{1+\epsilon}(1+\tilde{x}_1)^{2+\epsilon}}. \tag{3.57}$$

where we respectively applied the change of variable $x_2 = x_1\tilde{x}_2$ and $x_1 = x_2\tilde{x}_1$. As a result, we have transformed our integral with divergences appearing in the corner of the unit square by two integrals in which the divergence is concentrated in one of the variables. We can now expand the factorized divergence using Equation 3.52 to obtain

$$\int_0^1 dx_1 \int_0^1 dx_2 \frac{1}{(x_1+x_2)^{2+\epsilon}} = \frac{1}{\epsilon}\left(\int d\tilde{x}_2 \frac{1}{(1+\tilde{x}_2)^2} + \int d\tilde{x}_1 \frac{1}{(1+\tilde{x}_1)^2}\right) + \dots, \tag{3.58}$$

which is formulated in terms of convergent integrals that we can evaluate numerically using Monte Carlo integration.

The method of iterated sector decomposition generalizes this approach to an arbitrary number of Feynman parameters. To this end, one first builds N primary sectors from the integrand by singling

out one parameter in each sector:

$$\begin{aligned}
&\int_0^\infty \prod_i dx_i x_i^{n_i-1}\delta\left(1-\sum_i x_i\right)\frac{\mathcal{U}^{n-ld/2}}{\mathcal{F}^{n-(l+1)d/2}} \\
&=\sum_j \int_0^\infty dx_j x_j^{n_j-1}\int_0^{x_j}\prod_{i\neq j} dx_i x_i^{n_i-1}\delta\left(1-\sum_i x_i\right)\frac{\mathcal{U}^{n-ld/2}}{\mathcal{F}^{n-(l+1)d/2}} \\
&=\sum_j \int_0^\infty \frac{dx_j}{x_j}\int_0^1\prod_{i\neq j} d\tilde{x}_i \tilde{x}_i^{n_i-1}\delta\left(1-x_j\sum_{i\neq j}\tilde{x}_i\right)\frac{\mathcal{U}_j^{n-ld/2}(x_{i\neq j})}{\mathcal{F}_j^{n-(l+1)d/2}(x_{i\neq j})},
\end{aligned} \tag{3.59}$$

where in the last line we perfomed the change of variables $x_{i\neq j}\to x_i\tilde{x}_j$. The parameter x_j factorizes from the polynomials $\mathcal{U}$ and $\mathcal{F}$ because they are homogeneous and leave the first integral trivial:

$$\int_0^\infty \frac{dx_j}{x_j}\delta\left(1-x_j\sum_{i\neq j}\tilde{x}_i\right)=1. \tag{3.60}$$

For each primary sector j, one can then iteratively extract divergences by splitting the integration domain and rescaling the parameters. At each step, the original is expressed as a linear combination of integrals where k divergences have been factored out. Let us consider one such sector $I_{i_1,\dots,i_k}$ where the variables $i_1,\dots,i_k$ have been treated

$$I_{i_1,\dots,i_k}=\int_0^1\prod_{l=1}^k dx_{i_l}x_{i_l}^{a_{i_l}+\epsilon b_{i_l}}\int_0^1\prod_{\substack{m\notin\\ \{j,i_1,\dots,i_k\}}} dx_m x_m^{n_m-1}\frac{\mathcal{U}_{ji_1\dots i_k}^{n-ld/2}(x_m)}{\mathcal{F}_{ji_1\dots i_k}^{n-(l+1)d/2}(x_m)} \tag{3.61}$$

If there is a subset $\{x_m\}_{m\in S}$ such that $\mathcal{U}$ and $\mathcal{F}$ vanish when the parameters tend to 0, we split the integration over these parameters into subsectors as before. In each subsector one of the $x_{\hat{m}}$ is larger than the others and we use it to rescale all parameters and factor its divergence out, which generates new integrals $I_{i_1,\dots,i_k,\hat{m}}$ for each $\hat{m}\in S$. This produre is applied to all the integrals generated at each step and terminates when the polynomials cannot vanish, that is to say when

$$\mathcal{U}_{ji_1\dots i_k}=1+\dots \tag{3.62}$$

$$\mathcal{F}_{ji_1\dots i_k}=\sigma+\dots, \tag{3.63}$$

for some external scale σ. One then gets integrals of the form

$$\int_0^1\prod_{l=1}^k dx_{i_l}x_{i_l}^{a_{i_l}+\epsilon b_{i_l}}\int_0^1\prod_{\substack{m\notin\\ \{j,i_1,\dots,i_k\}}} dx_m x_m^{n_m-1}\frac{\mathcal{U}_{ji_1\dots i_k}^{n-ld/2}(x_m)}{\mathcal{F}_{ji_1\dots i_k}^{n-(l+1)d/2}(x_m)}, \tag{3.64}$$

in which we can insert the expansion in Equation 3.52 to generate many convergent integrals, which we can expand in ϵ, to obtain a Laurent series of convergent integrals in powers of ϵ and we can compute each integral using the Monte Carlo method.

Note that the method we have explained in this section only works in the Euclidean region, as otherwise $\mathcal{F}$ could vanish inside the integration region. Working in physical kinematics requires the integral

contour to be deformed around points where such singularities occur. We will not however discuss this aspect of numerical calculations, first of all because most of the numerical verifications we did of our results were performed in the Euclidean region, and because the discussion would be too long for this introductory exposition.

3.3 Multiple polylogarithms

Multiple polylogarithms (MPLs) form a class of functions that are essential to modern loop calculations. Most phenomenologically important results are formulated in terms of MPLs, such as the cross-section for Higgs boson production at NLO with full top mass dependence [51, 52], or that for the pair production of vector bosons at NNLO [139]. MPLs have been thoroughly investigated and have many well-known properties [159] which have lead to the development of a wealth of techniques to deal with them. In particular, they have a very rich algebraic structure that is essential to perform many calculations and this chapter will describe some aspects of this structure that have been relevant to deriving the results presented in this thesis.

MPLs are omnipresent in particle physics amplitudes and provide a very general class that ecompasses many generalizations of the logarithm such as classical polylogarithms or harmonic polylogarithms[160]. It has however been known for two decades that they are not the only functions that appear in loop integrals as one famously needs elliptic integrals to express the two-loop sunset integral [161] and the kite integral [162]. More than a mathematical curiosity, elliptic integrals have recently made their appearance in physically relevant calculations, such as that of the planar master integrals for the decay of a Higgs to three massless partons [163] and are expected to play a growing role in the field. Much less is known of the structure of the elliptic integrals that appear in loop integrals than we know about MPLs and we shall not discuss them in this section, as all integrals evaluated or used in this thesis can be expressed in terms of MPLs.

3.3.1 Definitions and basic properties

Multiple polylogarithms are defined as iterated integrals [164, 165]

$$G(a_1, \ldots, a_n; x) = \int_0^x \frac{dt}{t - a_1} G(a_2, \ldots, a_n; t), \tag{3.65}$$

where $G(; x) = 1$. The *entries* $a_1, \ldots, a_n$ are complex numbers and the variable x is also complex. The number of entries is called the *weight* of the MPL. A special case where the integral representation does not converge is that when all entries are 0:

$$G(\underbrace{0, \ldots, 0}_{n}; x) = \frac{1}{n!} \log^n(x) \tag{3.66}$$

The entries contain the information on the singularities of the polylogarithms. In particular, $G(a_1, \ldots, a_N; x)$ can have a divergence only at $x = a_1$ and branch cuts starting only at $x = a_i$. Some MPLs do reproduce this full structure, like $G(1, 2, x)$ which has a pole at $x = 1$, and branch cuts

starting at $x = 1$ and $x = 2$ as can be seen from its expression in terms of classical polylogarithms:

$$G(2, 1, x) = \mathrm{Li}_2(x - 1) + \log(1 - x)\log(2 - x) + \frac{\pi^2}{12}. \tag{3.67}$$

However, not all MPLs reproduce this phenomenon and some poles or branch cuts can be absent, as is the case for $G(0, 1, x) = -\mathrm{Li}_2(x)$, which has no divergence or branch cut starting at $x = 0$, but does have a branch cut starting at $x = 1$. It is in general necessary to study the functions weight by weight to understand their divergence structure.

As we hinted above, MPLs contain many classical functions that appear in loop calculations:

- Logarithms appear at the first level

$$G(a; x) = \log(1 - \frac{x}{a}) \tag{3.68}$$

- Powers of logarithms correspond to repeated arguments

$$G(\underbrace{a, \ldots, a}_{n}; x) = \frac{1}{n!}\log^n(1 - \frac{x}{a}), \tag{3.69}$$

  we shall see that this matches the definition of $G(0, \ldots, 0; x)$ in the framework of *shuffle regurlarization*.

- Classical polylogarithms, which can be defined as a series

$$\mathrm{Li}_n(x) = \sum_{k>1} \frac{x^k}{k^n} \tag{3.70}$$

  verify the property

$$\mathrm{Li}_n(x) = \int_0^x \frac{dt}{t}\mathrm{Li}_{n-1}(t) \tag{3.71}$$

  where $\mathrm{Li}_1(x) = -\log(1 - x)$. As a result,

$$G(\underbrace{0, \ldots, 0}_{n-1}, 1; x) = -\mathrm{Li}_n(x) \tag{3.72}$$

It is possible to generalize the series expansion of classical polylogarithms to multiple polylogarithms using nested sums. Let us define the following class of functions

$$\mathrm{Li}_{m_1,\ldots,m_k}(z_1, \ldots, z_k) = \sum_{0<n_1<\cdots<n_k} \frac{z_1^{n_1} \ldots z_k^{n_k}}{n_1^{m_1} \ldots n_k^{m_k}} \tag{3.73}$$

It is possible to express these functions as iterated integrals to see that

$$\mathrm{Li}_{m_1,\ldots,m_k}(z_1, \ldots, z_k) = (-1)^k G(\underbrace{0, \ldots, 0}_{m_k-1}, \frac{1}{z_k}, \ldots, \underbrace{0, \ldots, 0}_{m_1-1}, \frac{1}{z_1 \ldots z_k}; 1), \tag{3.74}$$

which provides a series representation for any G function.

One of the nice features of the MPLs over classical polylogarithms is that they form an algebra. To this end, let us first build a *shuffle algebra* over sequences of entries. We consider MPLs with entries taken from an *alphabet* Σ and build sequences, or *words* by concatenation. We consider a $\mathbb{Q}$-vector space $V(\Sigma)$ of linear combinations of letters[4]:

$$V(\Sigma) = \{q_1\sigma_1 + \dots q_n\sigma_n | q_i \in \mathbb{Q},\, \sigma_i \in \Sigma,\, n \in \mathbb{N}\}\,. \tag{3.75}$$

The space of words $W(\Sigma)$ is then the tensor algebra of $V(\Sigma)$ extended:

$$W(\Sigma) = \bigoplus_{n=0}^{\infty} V(\Sigma)^{\otimes n}, \tag{3.76}$$

and elementary tensors $w = \sigma_1 \otimes \dots \otimes \sigma_n$ represent words which we will write without the tensor product: $w = \sigma_1\sigma_2 \dots \sigma_n$. The factor $V(\Sigma)^{\otimes 0} = \mathbb{Q}1_\Sigma$, where 1_Σ is the empty word, is isomorphic to $\mathbb{Q}$. We can build an algebra based on the *shuffle product* $\sqcup\!\sqcup$ which is built iteratively:

Definition 6. *Let $\sigma, \sigma' \in V(\Sigma)$, the shuffle product of letters is defined by*

$$\sigma \sqcup\!\sqcup \sigma' = \sigma\sigma' + \sigma'\sigma, \tag{3.77}$$

Let $v, v' \in W(\Sigma)$ and $w = \sigma v, w' = \sigma' v'$, the shuffle product is defined iteratively on words by

$$w \sqcup\!\sqcup w' = \sigma(v \sqcup\!\sqcup w') + \sigma'(w \sqcup\!\sqcup v'). \tag{3.78}$$

The shuffle of two elementary words w, w' is the sum of all the ways the letters in w and w' can be interleaved while keeping the order in the letters of each word. For example:

$$ab \sqcup\!\sqcup cd = abcd + acbd + acdb + cabd + cadb + cdab \tag{3.79}$$

The reason we introduced this structure is that it is carried on to MPLs. Let us write MPLs whose entries form the word $w = \sigma_1 \dots \sigma_n$ as $G(\sigma_1, \dots, \sigma_n; x) = G_w(x)$ and extend this notation to $W(\Sigma)$ such that if $w = w_1 + w_2$, $G_w(x) = G_{w_1}(x) + G_{w_2}(x)$. This gives a very simple expression for the product of two MPLs:

$$G_{w_1}(x)G_{w_2}(x) = G_{w_1 \sqcup\!\sqcup w_2}(x). \tag{3.80}$$

[4]Explicitly, $V(\Sigma)$ can be constructed as the vector space of functions $\Sigma \to \mathbb{Q}$ with finite support, where a function f represents the linear combination $\sum_{\sigma\in\Sigma} f(\sigma)\sigma$

It is easy to see that this works explictly on MPLs with single entries:

$$
\begin{aligned}
G(\sigma_1;x)G(\sigma_2;x) &= \int_0^x dx_1 dx_2 \frac{1}{(x_1-\sigma_1)\,(x_2-\sigma_2)} && (3.81)\\
&= \int_0^x dx_1 \int_0^{x_1} dx_2 \frac{1}{(x_1-\sigma_1)\,(x_2-\sigma_2)} && (3.82)\\
&+ \int_0^x dx_2 \int_0^{x_2} dx_1 \frac{1}{(x_1-\sigma_1)\,(x_2-\sigma_2)} && (3.83)\\
&= G(\sigma_1,\sigma_2;x) + G(\sigma_2,\sigma_1;x) && (3.84)
\end{aligned}
$$

We can generalize this easily by recursion. Let us consider two elementary words $w = \sigma v$ and $w' = \sigma' v'$

$$
\begin{aligned}
G_w(x)G'_w(x) &= \int_0^x dt dt' \frac{G_v(t)G_{v'}(t')}{(t-\sigma)(t'-\sigma')}\\
&= \int_0^x dt \frac{1}{(t-\sigma)} \left(G_v(t) \int_0^t dt' \frac{G_{v'}(t')}{(t'-\sigma')} \right) + \int_0^x dt' \frac{1}{(t'-\sigma')} \left(G_{v'}(t') \int_0^{t'} dt \frac{G_v(t)}{(t-\sigma)} \right)\\
&= \int_0^x dt \frac{1}{(t-\sigma)} \left(G_v(t) G_{w'}(t) \right) + \int_0^x dt' \frac{1}{(t'-\sigma')} \left(G_{v'}(t') G_w(t') \right)\\
&= \int_0^x dt \frac{1}{(t-\sigma)} G_{v \sqcup\!\sqcup w'}(t) + \int_0^x dt' \frac{1}{(t'-\sigma')} G_{w \sqcup\!\sqcup v'}(t)\\
&= G_{\sigma(v \sqcup\!\sqcup w')}(x) + G_{\sigma'(w \sqcup\!\sqcup v')}(x) = G_{w \sqcup\!\sqcup w'}(x),
\end{aligned}
\qquad (3.85)
$$

where we identified the iterative definition of the shuffle product. One interesting thing is that, in the case of the words or the MPLs, the shuffle algebra defines a *graded* algebra: one can order the words by lengths and the MPLs by weight (which is just the entries length) and the shuffle product of elements of uniform weights n and m has a uniform weight $n + m$.

3.3.2 The Hopf algebra of polylogarithms

MPLs inherit more structure from $W(\Sigma)$ than its graded shuffle algebra. Indeed, we can build a Hopf algebra out of $W(\Sigma)$ that will propagate to the MPL. The Hopf algebra is a structure that will allow us to deconstruct complicated MPLs into tensor products of simpler functions, which provides us a tool to turn analytic problems such as the establishment of functional identities into algebraic problems.

Before defining what a Hopf algebra is, let us first reformulate the algebra structure in terms of tensor products to make the transition easier.

Theorem 2. *Let $\mathcal{A}$ be a unital algebra over a field $\mathbb{F}$ with product $\sqcup\!\sqcup$ and unit $1_{\mathcal{A}}$*

$$
\sqcup\!\sqcup \begin{cases} \mathcal{A} \times \mathcal{A} \to \mathcal{A} \\ (a,b) \to a \sqcup\!\sqcup b \end{cases} \qquad (3.86)
$$

There exists a linear map μ that reformulates the product as an operation on the tensor product $\mathcal{A} \otimes \mathcal{A}$ which

we can define on elementary tensors and extend by linearity:

$$\mu \begin{cases} \mathcal{A} \otimes \mathcal{A} \to \mathcal{A} \\ a \otimes b \to a \sqcup\!\sqcup b. \end{cases} \tag{3.87}$$

The associativity of the product implies the following property

$$\mu(\mu \otimes \mathrm{id}) = \mu(\mathrm{id} \otimes \mu), \tag{3.88}$$

where id is the identity, which can be nicely summarized by the following commutative diagram

$$\begin{array}{ccc} \mathcal{A} \otimes \mathcal{A} \otimes \mathcal{A} & \xrightarrow{\mu \otimes \mathrm{id}} & \mathcal{A} \otimes \mathcal{A} \\ \downarrow{\scriptstyle \mathrm{id} \otimes \mu} & & \downarrow{\scriptstyle \mu} \\ \mathcal{A} \otimes \mathcal{A} & \xrightarrow{\mu} & \mathcal{A} \end{array} \tag{3.89}$$

Because $\mathcal{A}$ contains a unit $1_{\mathcal{A}}$, there is a subalgebra isomorphic to the base field $\mathbb{F}$ (all elements of the form $\epsilon(f) = f1_{\mathcal{A}}$ for $f \in \mathbb{F}$). The existence of this isomorphism can be taken as an axiom instead of that of the existence of a unit to keep reformulating properties of the algebra as operations on tensor products:

Definition 7. *Let $\mathcal{A}$ be an algebra over $\mathbb{F}$, with a scalar product defined in the tensorial language*

$$s \begin{cases} \mathbb{F} \otimes \mathcal{A} \to \mathcal{A} \\ f \otimes a \to fa. \end{cases} \tag{3.90}$$

The algebra is unital if there exists a unit morphism $\epsilon : \mathbb{F} \to \mathcal{A}$ such that

$$\mu(\epsilon \otimes id) = \mu(id \otimes \epsilon) = s, \tag{3.91}$$

which we can again summarize in a diagram

$$\begin{array}{ccc} \mathbb{F} \otimes \mathcal{A} & \xrightarrow{\mathrm{id} \otimes \epsilon} & \mathcal{A} \otimes \mathcal{A} \\ \downarrow{\scriptstyle \epsilon \otimes \mathrm{id}} & \searrow^{s} & \downarrow{\scriptstyle \mu} \\ \mathcal{A} \otimes \mathcal{A} & \xrightarrow{\mu} & \mathcal{A} \end{array} \tag{3.92}$$

That this is equivalent to the existence of a unit is obvious: if such a map exists, $\epsilon(1_{\mathbb{F}})$ is the unit of $\mathcal{A}$ and if a unit $1_{\mathcal{A}}$ exists, the morphism $f \mapsto f1_{\mathcal{A}}$ fits the bill.

The idea of a Hopf algebra is to supplement the algebra structure with new operations that invert the arrows in the above diagrams. A Hopf algebra $\mathcal{H}$ is an algebra equipped with three new objects: a coproduct Δ, a counit η, and an antipode S.

The coproduct decomposes elements of the algebra as tensor products, working in the opposite sense to μ. It has to have two properties:

- Co-associativity:

$$(\Delta \otimes \mathrm{id})\Delta = (\mathrm{id} \otimes \Delta)\Delta \tag{3.93}$$

- Compatibility with the product

$$\Delta(\mu(a \otimes b)) = \Delta(a) \sqcup\!\sqcup \Delta(b), \tag{3.94}$$

where we define the multiplication of tensors as element-wise:

$$(a \otimes b) \sqcup\!\sqcup (a' \otimes b') = (a \sqcup\!\sqcup a') \otimes (b \sqcup\!\sqcup b'). \tag{3.95}$$

Indeed, the co-associativity results in the following diagram

$$\begin{array}{ccc} \mathcal{H}\otimes\mathcal{H}\otimes\mathcal{H} & \xleftarrow{\Delta\otimes\mathrm{id}} & \mathcal{H}\otimes\mathcal{H} \\ {\scriptstyle \mathrm{id}\otimes\Delta}\uparrow & & \uparrow{\scriptstyle \Delta} \\ \mathcal{H}\otimes\mathcal{H} & \xleftarrow{\Delta} & \mathcal{H} \end{array} \tag{3.96}$$

Note that one of the consequence of the co-associativity is that there is only one way to iterate the coproduct to generate tensors in $\mathcal{H}^{\otimes n}$:

$$\mathcal{H} \xrightarrow[\Delta]{} \mathcal{H}\otimes\mathcal{H} \xrightarrow[\Delta\otimes\mathrm{id}]{} \mathcal{H}\otimes\mathcal{H}\otimes\mathcal{H} \xrightarrow[\Delta\otimes\mathrm{id}\otimes\mathrm{id}]{} \dots \tag{3.97}$$

since all other choices are equivalent.

The counit η is a morphism $\mathcal{H} \to \mathbb{F}$ that operates the reverse of the unit operator ϵ, $\eta\epsilon = 1_{\mathbb{F}}$, and can be used to define $s^\dagger = (\eta \otimes \mathrm{id})\Delta$, which nicely reverse the diagram for the unit:

$$\begin{array}{ccc} \mathbb{F}\otimes\mathcal{H} & & \\ {\scriptstyle \eta\otimes\mathrm{id}}\uparrow & \nwarrow^{s^\dagger} & \\ \mathcal{H}\otimes\mathcal{H} & \xleftarrow{\Delta} & \mathcal{H} \end{array} \tag{3.98}$$

Finally, the antipode S is a new structure which does not have an equivalent in the algebra. It is an endomorphism $S : \mathcal{H} \to \mathcal{H}$ such that $\mu(S \otimes \mathrm{id})\Delta = \mu(\mathrm{id} \otimes S)\Delta = \epsilon\eta$, which is summarized as

$$\begin{array}{ccccc} \mathcal{H} & \xrightarrow{\eta} & \mathbb{F} & \xrightarrow{\epsilon} & \mathcal{H} \\ {\scriptstyle \Delta}\downarrow & & & & \uparrow{\scriptstyle \mu} \\ \mathcal{H}\otimes\mathcal{H} & & \xrightarrow[S\otimes\mathrm{id}]{\mathrm{id}\otimes S} & & \mathcal{H}\otimes\mathcal{H} \end{array} \tag{3.99}$$

The Hopf algebra of words Let us now build a Hopf algebra on the word algebra $W(\Sigma)$. We can construct a coproduct compatible with the shuffle product using the *deconcatenation* of words:

$$\Delta(\sigma_1 \dots \sigma_N) = \sum_{i=0}^{N} \sigma_1 \dots \sigma_i \otimes \sigma_{i+1} \dots \sigma_N, \tag{3.100}$$

where the two bounds of the sum are the products $1_\Sigma \otimes \sigma_1 \dots \sigma_N$ and $\sigma_1 \dots \sigma_N \otimes 1_\Sigma$. The fact that this coproduct is coassociative is easy to see: both ways of applying the coproduct a second time will yield the sum of all possible ways to cut the word into three pieces. This coproduct is compatible with the shuffle product, $\Delta(a \sqcup\!\sqcup b) = \Delta(a) \sqcup\!\sqcup \Delta(b)$, and, like the shuffle product, the coproduct respects the grading of $W(\Sigma)$ by the length

$$\Delta\left(\Sigma^n\right) \subset \bigoplus_{k=0}^{n} \left(\Sigma^k \otimes \Sigma^{(n-k)}\right). \tag{3.101}$$

The antipode is defined in the following way on elementary words

$$S(\sigma_1 \dots \sigma_N) = (-1)^N \sigma_N \dots \sigma_1, \tag{3.102}$$

and the counit η can be defined as the coefficient of the empty word. That is for a word $w = q_0 1_\Sigma + \sum_{i>0} w_i$, where $w_i \in V(\Sigma)^{\otimes i}$ are words of length i,

$$\eta(q_0 1_\Sigma + \sum_{i>0} w_i) = q_0. \tag{3.103}$$

A limited Hopf algebra of MPLs It seems easy to define a Hopf algebra on MPLs using the Hopf algebra of words. We have a morphism Φ which maps words to MPLs: $\Phi(a_1 \dots a_N) = G(a_1, \dots, a_N; x)$ and its inverse $\Phi^{-1}(G(a_1, \dots, a_N; x)) = a_1 \dots a_N$ and we can construct

$$\Delta_\Phi = (\Phi \otimes \Phi)\Delta\Phi^{-1}, \tag{3.104}$$

where Δ is the coproduct on words. For example, we can calculate

$$\begin{aligned} \Delta_\Phi G(a, b; x) &= (\Phi \otimes \Phi)\left(1 \otimes ab + a \otimes b + ab \otimes 1\right) \\ &= 1 \otimes G(a, b; x) + G(a; x) \otimes G(b; x) + G(a, b; x) \otimes 1. \end{aligned} \tag{3.105}$$

While this seems like a nice structure at first, it is very limited for several reasons. First of all, because there is a one-to-one mapping between words and the *integrands* of MPLs, this definition of the coproduct misses a significant amount of information related to the integration path. Indeed, MPLs have complicated branch cut structures and several equally valid MPLs can be defined for a given set of entries in some parts of $\mathbb{C}$. For example, on $\mathbb{R}^-$, $G(0, x)$ can be analytically continued as $G(0, -x) \pm i\pi$ but Δ_Φ will never know about this choice. This significantly limits the power of the coproduct to generate functional relations between MPLs. Furthermore, our definition fails when one considers the case of multiple variables or when one allows the entries to depend on x. In this latter case for example, non-trivial relations between MPLs appear that are not reflected in the algebra of words, see Appendix A for

details.

The Goncharov coproduct A better definition of a coproduct on MPLs has been constructed by Goncharov [166], which can deal with much more general cases as that inherited from the Hopf algebra of words. It is easier to define it using a notation for polylogarithms which allows us to choose the base point for the integration:

$$I(a_0; a_1, \dots a_n; a_{n+1}) = \int_{a_0}^{a_{n+1}} \frac{dt}{t - a_n} I(a_0; a_1, \dots a_{n-1}; t) \tag{3.106}$$

$$I(a_0; ; a_1) = 1. \tag{3.107}$$

This is in fact the same class of functions as the $G(a_1, \dots, a_n; x)$. Indeed, working iteratively, it is possible to express all I-MPLs into linear combinations of G-MPLs by cutting the integration at 0 so we can freely go back and forth between the two sets of integrals.

On these functions, the Goncharov coproduct is defined as

$$\begin{aligned}\Delta I(a_0; a_1, \dots a_n; a_{n+1}) = \sum_{k=0}^{n} \Bigg(& \sum_{0<i_1<\dots<i_k<n+1} \\ & I(a_0; a_{i_1}, \dots, a_{i_k}; a_{n+1}) \otimes \prod_{p=0}^{k} I(a_{i_p}; a_{i_p+1}, \dots, a_{i_{p+1}-1}; a_{i_{p+1}}) \Bigg)\end{aligned} \tag{3.108}$$

There is a nice graphical way of generating this formula:

- Draw a closed semi-circle, assign the values a_0 and a_{n+1} to the end-points of the diameter.
- Mark n points along the semi-circle and assign the values $a_1, \dots, a_n$ to them in a counterclockwise fashion.
- Each term is generated by selecting k points $a_{i_1}, \dots, a_{i_k}$ among the points on the semi-circle and drawing the convex polygon that has the vertices $a_0, a_{i_1}, \dots, a_{i_k}, a_{n+1}$:
  - The left-hand side of the tensor product is $I(a_0; a_{i_1}, \dots, a_{i_k}; a_{n+1})$
  - The terms of the product on the right-hand side are associated to the complementary convex polygons: the first is $I(a_0; a_1, \dots, a_{i_1-1}; a_{i_1})$ etc.

Let us look at how this works on the example of $I(a_0; a_1, a_2; a_3)$. As shown in Figure 3.3 we find

$$\begin{aligned}\Delta I(a_0; a_1, a_2; a_3) = {} & I(a_0; a_1, a_2; a_3) \otimes 1 + I(a_0; a_1; a_3) \otimes I(a_1; a_2; a_3) \\ & I(a_0; a_2; a_3) \otimes I(a_0; a_1; a_2) + 1 \otimes I(a_0; a_1, a_2; a_3)\end{aligned} \tag{3.109}$$

This definition is compatible with the product of MPLs and, as in the case of words, it is easy to see that it respects the grading of MPLs by the weight. It suffers, however from several issues when one considers specific choices for the entries.

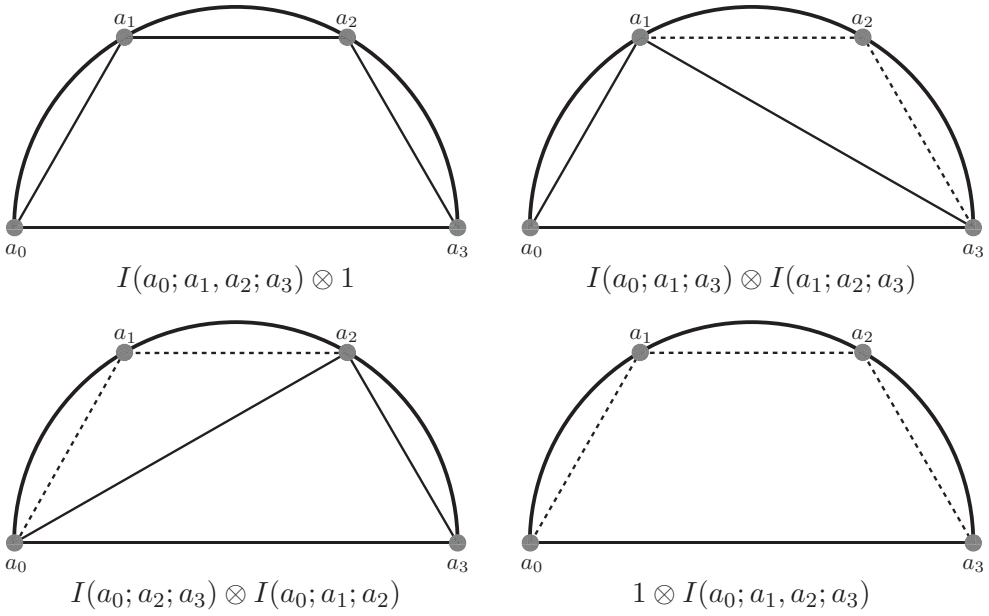


Figure 3.3: Graphical derivation of the coproduct $\Delta I(a_0;a_1,a_2;a_3)$.

The first issue is that divergences can appear in the coproduct of finite MPLs when several entries are identical:

$$\Delta I(1;0,1;x) = I(1,1,x)\otimes I(1;0;1) + \ldots, \tag{3.110}$$

And $I(1,1,x) = \int_1^x \frac{dt}{t-1}$ is a divergent integral. In order to be able to use the coproduct on any integral we encounter, we can regularize the MPLs in such a way that their algebraic structure is respected while removing all divergences. We use *shuffle regularization*, which is based in the structure of the MPLs. The basic idea is that one can use the shuffle product to factorize divergences of the form $G(a,\ldots,a;a)$ out of divergent MPLs systematically, which we can then set to 0. Let us consider for example the divergent MPL $G(-1,1,-1)$. We can consider the function $G(-1,1;x)$, which only has a divergence at $x=-1$. Using the shuffle algebra we can find

$$G(-1,1;x) = G(-1;x)G(1;x) - G(1,-1;x). \tag{3.111}$$

Now $G(1,-1;x)$ is perfectly regular at $x=-1$, while the other term has a factorized logarithmic divergence $G(-1;-1)$ and so we define $G^{\text{reg}}(-1,1;-1) = G(1,-1;-1)$. The reason why this approach works systematically is that the decomposition we operated in Equation 3.111 can be expressed in general as a *decomposition in Lyndon words*.

Definition 8. *Let Σ be an alphabet with an order $<$, extended to elementary words by the lexicographic order. A word $w=\sigma_1\ldots\sigma_n$ is said to be a Lyndon word if $w<\sigma_i\sigma_{i+1}\ldots\sigma_n$ for an$1<i\leq n$.*

For example $aabb$ and $aaabab$ are Lyndon words (for $a<b$), while baa is a counterexample. A theorem by Radford [167] shows that they provide a basis of the shuffle algebra of words:

Theorem 3. *Let Σ be an alphabet with an order $<$ extended to words by the lexicographic order. Let $L(\Sigma)_<$ be the set of Lyndon words. Any word $w\in W(\Sigma)$ can be expressed as a polynomial of elements of $L(\Sigma)_<$. In*

other words

$$\mathbb{Q}[L(\Sigma)_<] = W(\Sigma). \textit{ (with the shuffle product as product)} \tag{3.112}$$

We will use this to extract terms of the form $G(a,\dots,a;a)$. Indeed, if Σ admits a largest element S, then the only Lyndon word starting with S is S .

We can define the regularized MPLs G^{reg} in the following way. If $G(\sigma_1,\dots,\sigma_n;\sigma_{n+1})$ is divergent, $\sigma_1 = \sigma_{n+1}$. Let us define the alphabet $\Sigma = \sigma_1,\dots,\sigma_{n+1}$ (with redundant entries removed). We start by bulding an order $<_{\sigma_1}$ on Σ such that $\forall 1 < i \leq n+1$, $\sigma_i <_{\sigma_1} \sigma_1$, which is trivial because there is a finite number of letters. By the Radford theorem, the word $w = \sigma_1 \dots \sigma_n$ admits a decomposition into Lyndon words based on $<_{\sigma_1} L_{\sigma_1}$. The σ_1-regularized Lyndon word decomposition of w, $w^{\text{reg}}_{\sigma_1}$ is its Lyndon word decomposition with the singular word σ_1 set to 0. Using the duality between words and MPLs we define

$$G^{\text{reg}}(w;\sigma_{n+1}) = G_{w^{\text{reg}}_{\sigma_1}}(\sigma_{n+1}), \tag{3.113}$$

which reduces to $G(w;\sigma_{n+1})$ if it is not divergent, while if $\sigma_{n+1} = \sigma_1$, no elementary word in w^{reg} starts with σ_1 so that $G^{\text{reg}}(w;\sigma_{n+1})$ is regular. We can easily extend this definition to the I-MPLs as they also verify a shuffle algebra.

This regularization scheme preserves the algebraic structure of the MPLs. Indeed

$$G^{\text{reg}}_w(x)G^{\text{reg}}_{w'}(x) = G^{\text{reg}}_{w \sqcup\!\sqcup w'}(x), \tag{3.114}$$

and we can still build a coproduct on regulated integrals:

$$\Delta I^{\text{reg}}(a_0;\dots;a_{n+1}) = \sum_{0<i_1<\dots<i_k<n+1} I^{\text{reg}}(a_0;a_{i_1},\dots a_{i_k};a_{n+1}) \otimes \prod_{p=0}^{k} I^{\text{reg}}(a_{i_p};\dots;a_{i_{p+1}}). \tag{3.115}$$

Another issue that plagues some choices of entries appears when we consider ζ-values, $\zeta_n = \text{Li}_n(1)$. Indeed, the coproduct of classical polylogarithms is

$$\Delta\text{Li}_n(x) = \Delta I^{\text{reg}}(0;1,0,\dots,0;x) = 1 \otimes \text{Li}_n(x) + \sum_{k=0}^{n} \text{Li}_{n-k} \otimes \frac{\log^k(x)}{k!}. \tag{3.116}$$

Setting $x = 1$, we conclude that

$$\Delta\zeta_n = 1 \otimes \zeta_n + \zeta_n \otimes 1. \tag{3.117}$$

However, this implies that the coproduct is not in general compatible with multiplication. For example, we know that $\zeta_4 = \frac{2}{5}\zeta_2^2 = \pi^4/90$. If the coproduct was behaving as it should, we should have

$$\Delta\zeta_4 = \frac{2}{5}\left(\Delta\zeta_2\right)^2 = \zeta_4 \otimes 1 + 1 \otimes \zeta_4 + \frac{4}{5}\zeta_2 \otimes \zeta_2. \tag{3.118}$$

Among the possible solutions to restore Δ as a morphism, a very simple one is to work modulo ζ_2, that is modulo π^2, or even modulo π. A predecessor to the coproduct in loop calculations, the *symbol*, worked implicitly in this scheme [168], as did the original definition of the coproduct by Goncharov [166]. It is however better to keep as much information as possible in the coproduct and a consistent way of doing this is, in essence, to work modulo ζ_2 except in the leftmost term of the tensor product [169]. Let us call

the algebra of regulated MPLs $\mathcal{A}$, we can define the two-sided ideal generated by ζ_2, $\mathcal{I}^{\zeta_2} = \zeta_2\mathcal{A} + \mathcal{A}\zeta_2$ and the associated equivalence relation

$$x \sim y \iff x - y \in \mathcal{A}^{\zeta_2}. \tag{3.119}$$

We can now construct the *MPLs modulo* ζ_2, $\mathcal{A}^{\zeta_2}$ as the quotient $\mathcal{A}$ by this equivalence relation. If we now see Δ as an operation that maps $\mathcal{A}$ to $\mathcal{A} \otimes \mathcal{A}^{\zeta_2}$, we can consistently define

$$\Delta\zeta_{2n} = \zeta_{2n} \otimes 1, \tag{3.120}$$

which does not change the practical definition of Δ for the case of generic MPLs. By singling out the first entry of the coproduct, we also definitely fix how the iterated coproduct should work:

$$\mathcal{A} \xrightarrow[\Delta]{} \mathcal{A} \otimes \mathcal{A}^{\zeta_2} \xrightarrow[\Delta\otimes\mathrm{id}]{} \mathcal{A} \otimes \mathcal{A}^{\zeta_2} \otimes \mathcal{A}^{\zeta_2} \xrightarrow[\Delta\otimes\mathrm{id}\otimes\mathrm{id}]{} \dots \tag{3.121}$$

It was further conjectured in [168] that one can make this structure more powerful by working *modulo* π, that is to say to define

$$\Delta : \mathcal{A} \to \mathcal{A} \otimes \mathcal{A}^{\pi} \tag{3.122}$$

$$\Delta\pi \mapsto \pi \otimes 1. \tag{3.123}$$

Because $\zeta_2 \propto \pi^2$, this implies the definition of the coproduct on even ζ-values. We will systematically work with this definition of the coproduct.

Singling out the first entry has several consequences on the polylog. First of all, it makes it carry the information on the discontinuities of the function. Let us call $M_{x=a}$ the monodromy operator

$$M_{x=a}F(x) = \oint_a dx F(x), \tag{3.124}$$

where the integration is taken over a small contour around a, the following result is proven for MPLs with generic entries and conjectured to be general [168]:

$$\Delta M_{x=a}F(x) = (M_{x=a} \otimes \mathrm{id})\,\Delta F(x). \tag{3.125}$$

Similarly, derivatives only act on the rightmost entry

$$\Delta\frac{\partial F(x)}{\partial x} = \left(\mathrm{id} \otimes \frac{\partial}{\partial x}\right)\Delta F(x). \tag{3.126}$$

Both of these properties can be extended to the iterated coproduct.

We are now in position to use the coproduct to solve actual problems that arise in the calculation of loop integrals. In general, this structure is used to generate functional equations between different MPLs in a systematic and algorithmic way. In essence if we want to simplify a function $F(x)$, which is

expressed in terms of weight w MPLs, we can compute its coproduct

$$\Delta F(x) = 1 \otimes F(x) + F(x) \otimes 1 + \sum_i f_i(x) \otimes g_i(x). \tag{3.127}$$

Of course, the first two terms do not help us in any way, so it is usual to define $\Delta' F = \Delta F - 1 \otimes F(x) - F(x) \otimes 1$ to avoid carrying them in every calculation. On the other hand the rest of the coproduct is useful because it involves functions f_i, g_i that are of weight strictly less to w. We can therefore proceed in an interative manner, starting with weight 1 functions, *i.e.* ordinary logarithms, where functional equations are trivial and increase weight iteratively. If we want to provide a simplified expression $G(x) = F(x)$, we can use the lower weight identities to show that $\Delta' F(x) = \Delta' G(x)$. This, however, does not show that $F(x) = G(x)$, as Δ' has a non-trivial kernel. This kernel contains constants such as the powers of π and ζ-values. Indeed, for example

$$\Delta'(\mathrm{Li}_2(x) + \pi^2) = (1-x) \otimes x = \Delta'(\mathrm{Li}_2(x)). \tag{3.128}$$

Because the elements of the kernel of Δ' are usually rational multiples of known constants, it is however easy to fix them numerically. A downside to this way of generating functional equations is that one might generate a lot of useless relations at lower weights to generate a single useful equation and working constructively is often better.

Another approach to establishing functionnal identities is to work with the iterated coproduct to construct relations instead of generating tables of identities. Because the coproduct preserves the weight of objects we can uniquely decompose the coproduct on $\mathcal{A}_w$, the vector space of weight w MPLs:

$$\Delta\big|_{\mathcal{A}_w} = \sum_{k=0}^{w} \Delta_{k,w-k}, \tag{3.129}$$

where $\Delta_{k,w-k}$ maps to $\mathcal{A}_k \otimes \mathcal{A}_{w-k}$. Similarly we can decompose the iterated coproduct in pieces where each term in the tensor product has a definite weight and we will note them by extension $\Delta_{w_1,\ldots,w_n}$. Of particular interest is the term that has only weight one objects, $\Delta_{1,\ldots,1}$, which contains only logarithms. This term can be used to simplify arguments of MPLs: let us consider for example $\mathrm{Li}_2\left(\frac{1}{x+i\epsilon}\right)$ for $x \in [0,1]$

$$\begin{aligned}
\Delta_{1,1}\mathrm{Li}_2\left(\frac{1}{x+i\epsilon}\right) &= -\log\left(1-\frac{1}{x+i\epsilon}\right) \otimes \log\left(\frac{1}{x}\right) && (3.130)\\
&= \left(\log(1-x) - \log(-x-i\epsilon)\right) \otimes \log(x) && (3.131)\\
&= \log(1-x) \otimes \log(x) + \log(x) \otimes \log(x) + i\pi \otimes \log(x) && (3.132)\\
&= \Delta_{1,1}\left(-\mathrm{Li}_2(x) + \log^2(x) + i\pi\log(x)\right). && (3.133)
\end{aligned}$$

where we remembered the expression for $\Delta\mathrm{Li}_2$ from Equation 3.116. This, however, is not the final answer as there could be some irreducible weight two elements, *i.e.* some $c\pi^2$ term, and indeed by evaluating the two expression numerically one can see that they differ by $\pi^2/3$. For MPLs with higher

weight, using $\Delta_{1,\dots,1}$ does away with more information, which needs to be built iteratively. For example, at weight 3, $\pi^2 \log(x)$ is lost from $\Delta_{1,1,1}$, but can be reconstructed by taking $\Delta_{2,1}$. Using this approach, going from the terms of weight one and then reconstructing the missing pieces weight by weight, one can make sure to always deal with simple functions to construct functional relations.

3.3.3 Applications of the algebraic structure of polylogarithms

Analytic continuation of Harmonic Polylogarithms

A particular subclass of MPLs is of particular importance for the projects presented: harmonic polylogarithms. They are the MPLs whose entries are taken from the alphabet $\{0, 1, -1\}$ and have been studied extensively [160]. In particular, their analytic structure has been entirely elucidated:

- if the last entry is 0, the function may have a branch cut extending from 0 to $-\infty$,
- if the last entry is 1, the function may have a branch cut extending from 1 to $+\infty$,
- if one entry is -1, the function may have a branch cut extending from -1 to $-\infty$.

In this section, we will describe the technique proposed in [160] to analytically continuate harmonic polylogarithms to $-1 < x < 0$. In this case, a choice of branch cut is necessary for all functions whose rightmost entry is 0. The technique is closely related to the shuffle regularization and we can formulate it in an efficient algorithm using Lyndon words. The basic idea, like in shuffle regularization, is to factorize out functions of the form $G(0, \dots, 0; x)$ whose analytic behaviour around 0 is easy to describe, and leave only functions whose rightmost entry is not 0.

To this end, we use an order 0 on the alphabet such that 0 is the smallest element, for example $0 < 1 < -1$. Because of this Lyndon words cannot have 0 as their last entry because the word 0 is the smallest possible word.

Let us consider a harmonic polylogarithm $g(x) = G(\sigma_1, \dots, \sigma_k, 0, \dots 0)$, $\sigma_k \neq 0$ of weight w. In order to analytically continuate it to $-1 < x < 0$, we generate the set L_w of all Lyndon words of length smaller or equal to w and use them to build a basis $\left\{G_i^{(w)}\right\}$ of MPLs of weight w. There exists a set of rational numbers a_i such that

$$g(x) = \sum_i a_i G_i^{(w)}. \tag{3.134}$$

Using the shuffle algebra, we can express each product in the $G_I^{(w)}$ as a linear combination of MPLs over words of length w, which are linearly independent[5] [170]. As a result, we obtain an equation of the form

$$g(x) = \sum_{w \in \{0,\pm 1\}^4} b_w G(w; x), \tag{3.135}$$

where the b_w are linear combinations of the a_i. Because of the linear independence of the functions on the right-hand-side, this equation can only be solved if the following system is solved for the a_i

$$\begin{cases} b_{\sigma_1 \dots \sigma_k 0 \dots 0} = 1 \\ b_w = 0 \, \forall w \in \{0, \pm 1\}^4 \setminus \sigma_1 \dots \sigma_k 0 \dots 0. \end{cases} \tag{3.136}$$

[5] This is only the case for polylogarithms with constant entries.

Upon finding the a_i, we obtain an expression of $g(x)$ which involves only functions of the form $G(0,\ldots,0;x)$ and $G(\ldots,\pm1;x)$. We can now analytically continue the regular logarithms:

$$G(\underbrace{0,\ldots,0}_{n};x)\big|_{-1<x<0} = \frac{1}{n!}\left(\log(-x)\pm i\pi\right)^n, \tag{3.137}$$

and use the fact that for $-1<x<0$, $G(\sigma_1,\ldots,\sigma_n,\pm1,x) = G(-\sigma_1,\ldots,-\sigma_n,\mp1,-x)$.

Reduction to Canonical form

When evaluating Feynman integrals by direct integration, one takes repeated integrals over multi-variable rational functions, which most often yield MPLs. In such cases, one tries to transform the integral into a linear combination of terms of the form

$$\int_0^\infty dx \frac{G(a_1,\ldots,a_n;x)}{x-a_0} = \Big[G(a_0,a_1,\ldots,a_n;x)\Big]_0^\infty, \tag{3.138}$$

where all the entries a_j are independent of x. This kind of approach raises two issues:

- Apart from the simplest cases, it is rare that the a_j are independent of x.
- While the evaluation of the primitive at $x=0$ is easy (it is 0 unless all $a_j=0$, in which case there is a logarithmic divergence), taking the limit $x\to\infty$ is not trivial.

Both problems can be cast in a similar way for simple enough cases: given rational functions $A_j(x)$, can we in general express $G(A_1(x),\ldots,A_n(x);A_{n+1}(x))$ in terms of functions of the form $G(a_1,\ldots,a_k;x)$, where the a_j are independent of x? Explicitly, we want to find an expression of the form

$$G(A_1(x),\ldots,A_n(x);A_{n+1}(x)) = \sum_i G(a_1^{(i)},\ldots,a_n^{(i)};x) + \sum_i c_i^{(1)} G(a_1^{(i)},\ldots,a_{n-1}^{(i)};x) + \ldots, \tag{3.139}$$

where the $a_i^{(j)}$ and $c_i^{(k)}$ are independent of x and the $c_i^{(k)}$ are weight k constants (we limit ourselves to π, multiple ζ-values and MPLs evaluated at rational points). We will call such a form a *canonical form* of the MPL, introduced by [171]. In the following we outline the technique they devised to construct the canonical form, which is implemented in `PolyLogTools`.

We can consider what the iterated coproduct of such terms looks like

$$\Delta_{1,\ldots,1} c_i^{(k)} G(a_1^{(i)},\ldots,a_{n-k}^{(i)};x) = \left(\Delta_{1,\ldots,1} c_i^{(k)}\right) \sqcup\!\sqcup \left(\Delta_{1,\ldots,1} G(a_1^{(i)},\ldots,a_{n-k}^{(i)};x)\right) \tag{3.140}$$

$$= \left(\Delta_{1,\ldots,1} c_i^{(k)}\right) \sqcup\!\sqcup \Big(\log(x-a_{n-k}^{(i)})\otimes\cdots\otimes\log(x-a_1^{(i)}) \tag{3.141}$$

$$+\text{non maximal tensors,}\Big)$$

where we define maximal tensors as tensors in which no entry is independent of x. As a result, each term with a weight k constant has exactly one tensor with $n-k$ entries that depend on x [171]. If a function $g(x) = G(A_1(x),\ldots,A_n(x);A_{n+1}(x))$ admits a canonical form, we can use the iterated coproduct to fix this form up to some constants $c_i^{(k)}$ if its iterated coproduct can be expressed in terms of degree one

polynomials. Indeed let us consider the maximal tensors of the iterated coproduct $T = \Delta_{1,\dots,1} g(x)$:

$$T = \sum_i \log(a_n^{(i)} x + b_n^{(i)}) \otimes \cdots \otimes \log(a_1^{(i)} x + b_1^{(i)}) + \text{non maximal tensors.} \tag{3.142}$$

Let us construct the linear map:

$$\phi_x^{(1)} : \log(a_n^{(i)} x + b_n^{(i)}) \otimes \cdots \otimes \log(a_1^{(i)} x + b_1^{(i)}) \to \begin{cases} G(-\frac{b_n}{a_n}, \dots, -\frac{b_1}{a_1}; x) \text{ if all } a_i \neq 0 \\ 0 \text{ otherwise,} \end{cases} \tag{3.143}$$

which maps each maximal tensor to a canonical form MPL with a matching maximal tensor. We can then build

$$T_1 = T - \Delta_{1,\dots,1} \phi_x^{(1)}(T), \tag{3.144}$$

which only contains non maximal tensors. Let us look at next-to-maximal tensors in T_1, which by Equation 3.141 must come from functions of the form $c^{(1)} G(a_1, \dots, a_{n-1}; x)$. There are few possible weight-one functions: they will be either π or some $\log(\chi)$ for some number χ. In the case when $c = \pi$, a single tensor is generated:

$$\Delta_{1,\dots,1} \pi G(a_1, \dots, a_{n-1}; x) = \pi \otimes \Delta_{1,\dots,1} G(a_1, \dots, a_{n-1}; x), \tag{3.145}$$

while logarithms will be shuffled to every position:

$$\Delta_{1,\dots,1} \log(\chi) G(a_1, \dots, a_{n-1}; x) = \log(\chi) \sqcup\!\sqcup \Delta_{1,\dots,1} G(a_1, \dots, a_{n-1}; x) \tag{3.146}$$

$$= \log(\chi) \otimes \Delta_{1,\dots,1} G(a_1, \dots, a_{n-1}; x) + \dots \tag{3.147}$$

As a result, to fix these functions, it is always sufficient to consider tensors t such that

$$t = c \otimes (a_1 x + b_1) \otimes (a_{n-1} x + b_{n-1}). \tag{3.148}$$

Indeed, we can build the map

$$\phi_x^{(2)} : \begin{cases} \log(c) \otimes \log(a_1 x + b_1) \otimes \cdots \otimes \log(a_{n-1} x + b_{n-1}) \mapsto \\ \qquad\qquad \log(c) \phi_x^{(1)} \left(\log(a_1 x + b_1) \otimes \cdots \otimes \log(a_{n-1} x + b_{n-1}) \right) \\ \text{all other tensors} \mapsto 0. \end{cases} \tag{3.149}$$

and use it to eliminate all next-to-minimal tensors from the symbol of our function:

$$T_2 = T_1 - \Delta_{1,\dots,1} \phi_x^{(2)}(T_1). \tag{3.150}$$

In physically relevant situations, these two steps are more than enough and one usually finds that T_1 or T_2 is zero, therefore stopping the iteration. In general, this does however not fix the function entirely, as we must remember that the iterated coproduct has a non-zero kernel which contains all the ζ values, which we must fix by taking other parts of the iterated coproduct. Let us for example consider the

function $\mathrm{Li}_3\left(\frac{1}{x+i\epsilon}\right)$ for $0<x<1$:

$$\Delta_{1,1,1}\mathrm{Li}_3\left(\frac{1}{x}\right) = -\log\left(\frac{x-1}{x}\right)\otimes\log\left(\frac{1}{x}\right)\otimes\log\left(\frac{1}{x}\right) \tag{3.151}$$

$$= -\left(\log(1-x)+i\pi\right)\otimes\log(x)\otimes\log(x)+\log(x)\otimes\log(x)\otimes\log(x) \tag{3.152}$$

$$\tag{3.153}$$

We can use our maps $\phi_x^{(1)}$ and $\phi_x^{(2)}$ to find that

$$\Delta_{1,1,1}\mathrm{Li}_3\left(\frac{1}{x}\right) = \Delta_{1,1,1}\left(-G(0,0,1;x)-i\pi G(0,0;x)+G(0,0,0;x)\right) \tag{3.154}$$

and indeed, the recursion ends there. However, if we apply $\Delta_{2,1}$ to the difference of these two functions, we find that there is a leftover $-\frac{1}{3}\pi^2\otimes\log(x)$. We can build further maps in general to deal with such terms and their higher weight equivalents, but in this case, we can trivally find that

$$\Delta_{2,1}\mathrm{Li}_3\left(\frac{1}{x}\right) = \Delta_{2,1}\left(-G(0,0,1;x)-i\pi G(0,0;x)+G(0,0,0;x)-\frac{\pi^2}{3}G(0;x)\right). \tag{3.155}$$

By catching ever higher weight constants in the leftmost term of the iterated coproduct, we can fix the canonical form of functions up to irreducible constants. These can be obtained numerically by evaluating the functions for enough values of x, as we did in the case of $\mathrm{Li}_2(1/x)$. This generates a linear system with rational solutions, which can be found very efficiently using the PSLQ algorithm [172]. In the particular case of $\mathrm{Li}_3(1/x)$, no such constant arises and indeed

$$\Delta_{2,1}\mathrm{Li}_3\left(\frac{1}{x}\right) = -G(0,0,1;x)-i\pi G(0,0;x)+G(0,0,0;x)-\frac{\pi^2}{3}G(0;x). \tag{3.156}$$

Note that not all constants are in the kernel of the iterated coproduct. For example, $\Delta_{1,1}\mathrm{Li}_2\left(1/2\right) = -\log(2)\otimes\log(2)$. However, this example shows that there is a high degree of degeneracy among constants in $\Delta_{1,\ldots,1}$ and it is therefore a a better choice to leave them aside and fix all the $c^{(k)}$ constants of the canonical expression using $\Delta_{k,1,\ldots,1}$ since all such constants appear explicitly in the leftmost entry.

As this technique is used to compute integrals over many Feynman parameter, it is by design very easily generalizable to the multivariate case [171]. Let us consider N variables $x_1,\ldots,x_N$ with this order. A function is said to be in a canonical form with respect to this ordering if it is a linear combination of functions

$$G(\vec{a}_1;x_1)G(\vec{a}_2;x_2)\ldots G(\vec{a}_N;x_N), \tag{3.157}$$

such that $\vec{a}_i$ is independent of the x_j for $j\leq i$. A sufficient condition for the algorithm to work is that the symbol of the function considered be of degree one with respect to all the x_i. Let us build maps that

work recursively to construct these products:

$$\phi_{x_k x_{k+1} \dots x_N} : \begin{cases} \log(a_1 x_k + b_1) \otimes \cdots \otimes \log(a_n x_k + b_n) \mapsto G(-\frac{b_1}{a_1}, \dots, -\frac{b_n}{a_n}; x_k) \\ \vdots \\ \log(b_1) \otimes \cdots \otimes \log(b_p) \otimes \log(a_{p+1} x_k + b_{p+1}) \otimes \log(a_n x_k + b_n) \mapsto \\ \qquad \phi_{x_{k+1} \dots x_N} \left(\log(b_1) \otimes \dots \log(b_p) \right) G(-\dfrac{b_{p+1}}{a_{p+1}}, \dots, -\dfrac{b_n}{a_n}; x_k) \\ \vdots \end{cases} \tag{3.158}$$

where the a_i, b_i are independent of x_j for $j \leq i$. Indeed, $\phi_{x_1 \dots x_N}$ will match maximal tensors to MPLs in canonical form with respect to x_1 and non-maximal tensors as products involving only maximal tensors. If the functions admits a canonical form, this removes all tensors that depend on at least one variable in each entry. We can deal with constants as in the single variable case, by using $\Delta_{k,1,\dots,1}$, which will catch order k constants in its leftmost entry.

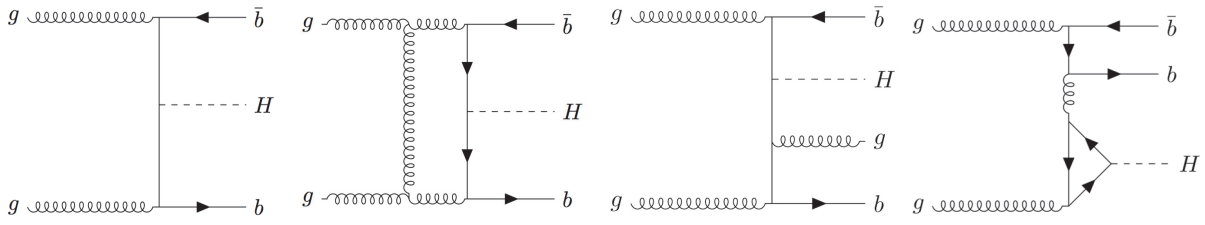
A critical point in this algorithm is that indeed the iterated coproduct be expressed in terms of linear functions of the variables. While this does not cover all MPLs that appear when integrating loop integrals, we could apply it to all relevant situations in this thesis to simplify expressions, which allowed us to perform integrals explicitly.

4

Matching the decay $H \to b\bar{b}$ between the Standard Model and the Higgs Effective Field Theory

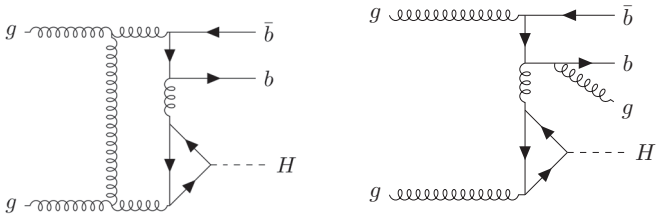
THE BOTTOM QUARK plays an interesting role in Higgs physics. Despite having a relatively low coupling ($y_b \simeq 0.025$), the decay $H \to b\bar{b}$ is the dominant channel for a SM Higgs boson with a $125\ \mathrm{GeV}$ mass due to kinematics and phase space effects. Observing this decay is however challenging because of the large backgrounds generated by QCD, especially in the gluon-fusion production mode [7], and has for now only been probed in weak production modes: vector boson fusion [173, 174] and associated production [175, 176], which is the most sensitive channel, yielding a signal strength for the decay $\mu_{bb} = 0.7 \pm 0.3$. Instead of studying this decay, the interaction of the Higgs boson with bottom quarks can be tested using production mechanisms in which it plays a role. The main such process is gluon fusion, where around 5% of the total cross section is contributed by bottom quarks [177–180], meaning that precise measurements of this cross section can put constraints on $Hb\bar{b}$ couplings. Another possible avenue is to study the associated production of the Higgs boson and bottom quarks, which has a comparable total production cross section to associated production with top quarks (around $0.5\ \mathrm{pb}$ at $13\ \mathrm{TeV}$) making it a minor production mode in the SM [181]. Furthermore, selection cuts that could isolate this mode would drastically reduce the cross section: requiring two b-jets harder than $20\ \mathrm{GeV}$ yields an efficiency of less than 5% [44, 182]. While a direct observation is not yet possible with the current data, it is still a useful handle to constrain modified Higgs sectors that could make this process more important, such as the MSSM with a sizeable $\tan\beta$ [181, 183].

The cross section for $b\bar{b}H$ associated production has been calculated to NLO QCD accuracy by [44]. At LO in QCD ($\mathcal{O}(\alpha_s^2)$), the production is mediated by the bottom Yukawa coupling, as shown in Figure 4.1a, while at NLO ($\mathcal{O}(\alpha_s^3)$), contributions from both the bottom and top Yukawa couplings appear, as shown in Figures 4.1b-4.1d, both of which have the same numerical impact on the prediction,



(a) LO - $\mathcal{A}_b^{(0)}$ **(b)** NLO - $\mathcal{A}_b^{(1V)}$ **(c)** NLO - $\mathcal{A}_b^{(1R)}$ **(d)** NLO - $\mathcal{A}_t^{(0)}$

Figure 4.1: Examples of Feynman diagrams for $b\bar{b}H$ production at LO and at NLO, which contain virtual and real bottom-induced diagrams and Born-level top-induced diagrams. The corresponding amplitudes are named $\mathcal{A}_b^{(0)}$, $\mathcal{A}_b^{(1V)}$, $\mathcal{A}_b^{(1R)}$ and $\mathcal{A}_t^{(0)}$.



(a) $\mathcal{A}_t^{(1V)}$ **(b)** $\mathcal{A}_t^{(1R)}$

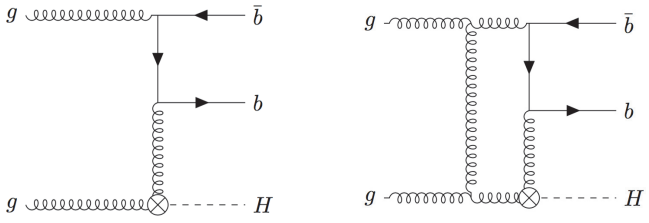
Figure 4.2: Virtual $\left(\mathcal{A}_t^{(1V)}\right)$ and real emission $\left(\mathcal{A}_t^{(1R)}\right)$ diagrams contributing to associated $b\bar{b}H$ production at $\mathcal{O}\left(y_t y_b \alpha_s^4\right)$ through their interference with $\mathcal{A}_b^{(0)}$ and $\mathcal{A}_b^{(1R)}$.

as shown in Table 4.1. While NLO corrections improve the scale dependence of the prediction, the

	LO	NLO (y_b^2)	NLO $(y_b^2 + y_b y_t)$
σ	$0.478^{+60\%}_{-35\%}$	$0.448^{+20\%}_{-20\%}$	$0.411^{+25\%}_{-28\%}$
$\delta\sigma/\sigma_{\text{LO}}$	–	6%	14%

Table 4.1: Cross sections for the associated production $b\bar{b}H$ at LO and NLO, including and excluding top quark effects.

contribution generated by the top Yukawa coupling itself suffers from a large scale dependence, which makes the full NLO QCD prediction more scale dependent than when excluding top quark effects. It would therefore be desirable to improve the prediction for the top Yukawa coupling dependence of the cross section by deriving and extra order in QCD, *i.e.* calculating the $\mathcal{O}\left(y_t y_b \alpha_s^4\right)$ part of the cross section, which is generated by the interference of the diagrams shown in Figure 4.2b with the LO diagrams of Figure 4.1a, as well as by real emission intereference of the diagrams shown in Figure 4.2b with those in Figure 4.1c. In order to make it easier to follow the discussion, let us write down the elementary cross



(a) $\hat{\mathcal{A}}_t^{(0)}$ **(b)** $\hat{\mathcal{A}}_t^{(1V)}$

Figure 4.3: Examples of Born-level and virtual correction diagrams for top-induced $b\bar{b}H$ production in the infinite top mass limit.

section explicitly

$$\begin{aligned} d\sigma \propto \alpha_s^2 y_b^2 \left|\mathcal{A}_b^{(0)}\right|^2 + \alpha_s^3 \left[y_b^2 \left(2\text{Re}\left(\mathcal{A}_b^{(0)} \mathcal{A}_b^{(1V)}\right) + \int d\Phi \left|\mathcal{A}_b^{(1R)}\right|^2 \right) + 2 y_b y_t \text{Re}\left(\mathcal{A}_b^{(0)} \mathcal{A}_t^{(0)}\right) \right] \\ + 2\alpha_s^4 y_t y_b \text{Re}\left(\mathcal{A}_b^{(0)} \mathcal{A}_t^{(1V)} + \mathcal{A}_b^{(1V)} \mathcal{A}_t^{(0)} + \int d\Phi \mathcal{A}_b^{(1R)} \mathcal{A}_t^{(1R)} \right), \end{aligned} \tag{4.1}$$

where the amplitudes have been defined in Figures 4.1 and 4.2 and the integrals are taken over the real emission phase space. The first line of Equation 4.1 is what has been calculated in [44]. While an exact calculation of the $\mathcal{O}\left(y_t y_b \alpha_s^4\right)$ contribution is still out of reach at the moment due to the complicated two loop calculation with four external particles and three different masses required for the evaluation of $\mathcal{A}_t^{(1V)}$, the calculation can be performed in the HEFT, in which all top loops would be reduced to contact interactions, making the most complicated amplitude a one-loop calculation, as shown in Figure 4.3. However, this calculation cannot readily be performed as the infinite top mass limit of the HEFT also yields power suppressed corrections to SM parameters and in particular, as we shall soon see, the bottom quark Yukawa coupling is corrected in the following way:

$$y_b^{\text{HEFT}} = y_b + \frac{y_t \alpha_s^2}{m_t} \delta y_b, \tag{4.2}$$

and these corrections generate terms of order $\mathcal{O}\left(y_t y_b \alpha_s^4\right)$ through the squares of the diagrams shown in Figure 4.1a. As a result, in the HEFT, the elementary cross section can be decomposed as

$$\begin{aligned} d\sigma^{\text{HEFT}} \propto \alpha_s^2 y_b^2 \left|\mathcal{A}_b^{(0)}\right|^2 + \alpha_s^3 \left[y_b^2 \left(2\text{Re}\left(\mathcal{A}_b^{(0)} \hat{\mathcal{A}}_b^{(1V)}\right) + \int d\Phi \left|\mathcal{A}_b^{(1R)}\right|^2 \right) + 2 y_b y_t \text{Re}\left(\mathcal{A}_b^{(0)} \hat{\mathcal{A}}_t^{(0)}\right) \right] \\ + \alpha_s^4 y_t y_b \left(2\text{Re}\left(\mathcal{A}_b^{(0)} \hat{\mathcal{A}}_t^{(1V)} + \mathcal{A}_b^{(1V)} \hat{\mathcal{A}}_t^{(0)} + \int d\Phi \mathcal{A}_b^{(1R)} \hat{\mathcal{A}}_t^{(1R)} \right) + \frac{\delta y_b}{m_t} \left|\mathcal{A}_b^{(0)}\right|^2 \right), \end{aligned} \tag{4.3}$$

where, in contributions with a hat, top quark loops have been replaced by the HEFT contact interaction. In this cross section, the only contribution that cannot be readily calculated using automated tools is the power suppressed bottom Yukawa correction. We derive this correction in this chapter, and thus provide the last missing piece necessary for the improvement of the associated $b\bar{b}H$ production cross section prediction.

One of the advantages of using EFTs to make approximations is that they are universal. That is to say that they have a finite set of parameters in their Lagrangian (at each order) that can be determined once and for all and be used for arbitrarily many predictions. We can make use of this universality to extract the value of the parameter we are interested in by matching its contribution to the simplest observable possible with the prediction of the SM. The easiest process we can use to this end happens to be a scalar form factor contributing to the decay $H \to b\bar{b}$. We will calculate in each theory

$$\mathcal{A} = \sum_{\sigma\sigma' ij} \bar{v}_{\bar{b}}^{\sigma'}(p_2) u_b^{\sigma}(p_1) \delta^{ij} \mathcal{A}_{H\to b\bar{b}} = C_A \mathrm{Tr}\left((\not{p}_1 + m_b) \mathcal{A}^{\mathrm{amp}}_{H\to b\bar{b}} (\not{p}_2 - m_b) \right), \tag{4.4}$$

where $u(p_1)$ ($\bar{v}(p_2)$) is the conjugate of the wavefunction of the (anti) b-quark, the sums go over their colors and helicities, $\mathcal{A}_{H\to b\bar{b}}$ is the complete amplitude for the decay $H \to b\bar{b}$, and

$$\mathcal{A}_{H\to b\bar{b}} = \bar{u}_b^{\sigma}(p_1) \delta_{ij} \mathcal{A}^{\mathrm{amp}}_{H\to b\bar{b}} v_{\bar{b}}^{\sigma'}(p_2). \tag{4.5}$$

In order to obtain a SM Feynman diagram for this process that has power-suppressed terms, we need to go to the two-loop level, as shown in Fig. 4.4. We need to compute the $1/m_t$ expansion of these diagrams and match them to the relevant HEFT diagrams in order to extract the desired correction to the bottom Yukawa.

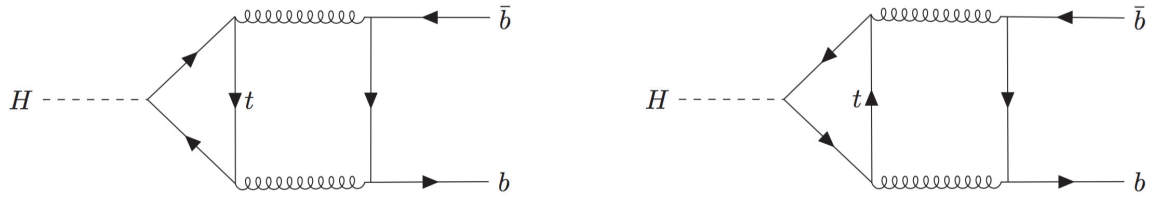


Figure 4.4: The two $\mathcal{O}\left(y_t g_s^4\right)$ diagrams that contribute to $H \to b\bar{b}$

In the first section of this chapter, we will perform the easier HEFT calculation, which involves both a tree-level and a one-loop diagram. Section 4.2.3 will then be dedicated to the calculation of the corresponding two-loop process in the SM in a $1/m_t$ expansion and in Section 4.3, we will compare the two results to extract the first power-suppressed correction to the bottom Yukawa in the HEFT.

4.1 Higgs-to-bottom decay in the Higgs Effective Field Theory

In the HEFT, two diagrams will be necessary to match that in Figure 4.4. The first one is shown in Figure 4.5a and is the contribution that intuitively comes to mind when one thinks of integrating out top quarks from the SM diagrams. This one-loop diagram is divergent and its pole will be cancelled through the renormalization of the bottom Yukawa, which appears in the tree-level diagram shown in Figure 4.5b. This already shows that it receives power-suppressed corrections of order $\mathcal{O}\left(y_t \alpha_s^2\right)$, but the HEFT calculation alone will only allow us to find the divergent piece of this correction. The finite part of this correction, which we will obtain at the end of this chapter, can only be extracted from the

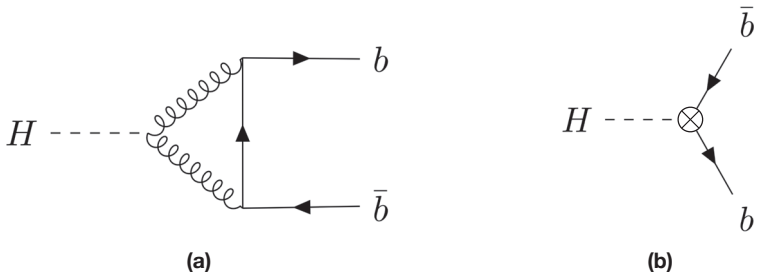


Figure 4.5: The two Feynman diagrams that contribute to $H \to b\bar{b}$ in the HEFT. We denote the HEFT bottom Yukawa with a crossed dot to highlight the fact that it is not equal to the SM Yukawa.

matching with the SM.

Another way to understand the presence of these two contributions will appear in the next section: the most complicated of the loop integrals that appear in the SM two-loop amplitude will have two regions in the $1/m_t$ expansion, which correspond to routing hard momenta through the top loop only or through the complete diagram. These respectively have a diagramatical interpretation as shrinking the top loop in Figure 4.4 to obtain Figure 4.5a and shrinking the two loops to obtain Figure 4.5b.

Since we wish to extract the power-suppressed corrections to a coupling that is present in the SM, it is important that our notation makes a proper separation between the couplings in the SM and the HEFT to avoid confusion. The interactions relevant for our calculation are shown in Table 4.2

Coupling	Value	Feynman Rule
$\mathcal{L}_{y_b} = -\dfrac{C_{y_b}}{\sqrt{2}} h\bar{b}b$	$C_{yb} = y_b + \mathcal{O}\left(\dfrac{1}{m_t}\right)$	$-i\dfrac{C_{y_b}}{\sqrt{2}}$
$\mathcal{L}_{Hgg} = -\dfrac{1}{4} C_{Hgg} G_{\mu\nu} G^{\mu\nu} h$	$C_{Hgg} = \dfrac{\alpha_s}{3\pi v}$	$iC_{Hgg}\left(g^{\mu\nu} p_1 \cdot p_2 - p_1^\nu p_2^\mu\right)$
$\mathcal{L}_{bG} = g_s G_a^\mu \bar{b} T^a \gamma_\mu b$	Fixed by gauge invariance	$ig_s \gamma^\mu T^a$

Table 4.2: HEFT interactions relevant for the calculation of the $\mathcal{O}(g_s^4 y_t)$ contribution to $H \to \bar{b}b$ and their leading-order values in the $1/m_t$ expansion.

The part of $\mathcal{A}$ coming from Fig 4.5b is trivial to evaluate;

$$\mathcal{A}_b = -i\frac{C_{y_b} C_A}{\sqrt{2}}\left(s - 4m_b^2\right). \tag{4.6}$$

We can now move on to the calculation of the one-loop diagram $\mathcal{A}_{\text{gf}}$. Reading the Feynman rules, the expression for this diagram in dimensional regularization is

$$\mathcal{A}_{\text{gf}} = iC_{Hgg}\left(ig_s\right)^2 \int \frac{d^d k}{(2\pi)^d} \frac{\text{Tr}\left((-i)(\not{p}_1 + m_b)\gamma_\mu i(\not{k} + m_b)\gamma_\nu (-i)(\not{p}_2 - m_b)\right)}{(k+p_2)^2\left(k^2 - m_b^2\right)(k-p_1)^2}\left(g^{\mu\nu} p_1 \cdot p_2 - p_1^\nu p_2^\mu\right). \tag{4.7}$$

Until Subsection 4.1, all parameters are understood to be bare, d-dimensional quantities and we will

not use super/subscripts to separate them from renormalized parameters.

While $\mathcal{A}_{\text{gf}}$ is a relatively simple amplitude, it is a good opportunity to illustrate the modern approach to loop calculations. We will describe how we matched the scalar integrals of our amplitude to a topology and reduced them to master integrals, which we will then evaluate.

Since there is only one diagram, choosing the topology is easy:

$$\mathrm{I}\left(n_1,n_2,n_3\right) = \int \frac{d^dk}{(2\pi)^d}\frac{1}{D_1^{n_1}D_2^{n_2}D_3^{n_3}} = \int \frac{d^dk}{(2\pi)^d}\frac{1}{\left(\left(k-p_1\right)^2\right)^{n_1}\left(\left(k+p_2\right)^2\right)^{n_2}\left(k^2-m_b^2\right)^{n_3}} \tag{4.8}$$

Which is summed up in Figure 4.6.

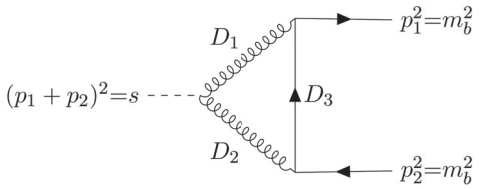


Figure 4.6: Graphical definition of the topology, defined as the set of denominators appearing in this diagram.

Reducing the integrals with numerators can be done using the following rules

$$k\cdot k = D_3 + m_b^2, \qquad k\cdot p_1 = \frac{D_3-D_1}{2} + m_b^2, \qquad k\cdot p_2 = \frac{D_2-D_3}{2} - m_b^2. \tag{4.9}$$

In practice, the generation of the amplitude and its reduction was handled by the chain of tools detailed in Appendix B. The diagrams were generated with `QGRAF` and translated with `qgraf-xml-parser` into `FORM` input files to perform the index contractions and Dirac algebra traces. These contracted expressions were then imported in `Mathematica` for further treatment. The integrals were matched to the defined topology using `SAIF`.

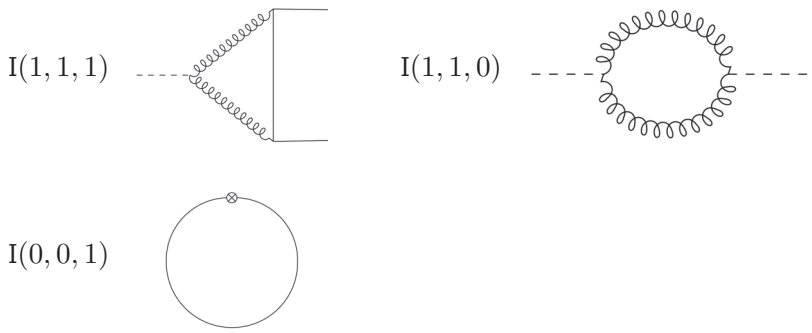
The result of these operations yields:

$$\begin{aligned}
\mathcal{A}_{\text{gf}} =& -4ds^2\,\mathrm{I}\left(1,1,1\right)m_bC_{Hgg}\,g_s^2 - 4d\,\mathrm{I}\left(-1,1,1\right)m_bC_{Hgg}\,g_s^2 + 8d\,\mathrm{I}\left(0,0,1\right)m_bC_{Hgg}\,g_s^2\\
&- 8d\,\mathrm{I}\left(0,1,0\right)m_bC_{Hgg}\,g_s^2 + 8ds\,\mathrm{I}\left(0,1,1\right)m_bC_{Hgg}\,g_s^2 + 8d\,\mathrm{I}\left(1,0,0\right)m_bC_{Hgg}\,g_s^2\\
&- 8ds\,\mathrm{I}\left(1,0,1\right)m_bC_{Hgg}\,g_s^2 - 4d\,\mathrm{I}\left(1,1,-1\right)m_bC_{Hgg}\,g_s^2 + 8ds\,\mathrm{I}\left(1,1,0\right)m_bC_{Hgg}\,g_s^2\\
&+ 8s^2\,\mathrm{I}\left(1,1,1\right)m_bC_{Hgg}\,g_s^2 - 32\,\mathrm{I}\left(0,1,1\right)m_b^3C_{Hgg}\,g_s^2 - 32\,\mathrm{I}\left(1,1,0\right)m_b^3C_{Hgg}\,g_s^2\\
&+ 32s\,\mathrm{I}\left(1,1,1\right)m_b^3C_{Hgg}\,g_s^2 + 8\,\mathrm{I}\left(-1,1,1\right)m_bC_{Hgg}\,g_s^2 - 24\,\mathrm{I}\left(0,0,1\right)m_bC_{Hgg}\,g_s^2\\
&+ 16\,\mathrm{I}\left(0,1,0\right)m_bC_{Hgg}\,g_s^2 - 16s\,\mathrm{I}\left(0,1,1\right)m_bC_{Hgg}\,g_s^2 + 16\,\mathrm{I}\left(1,-1,1\right)m_bC_{Hgg}\,g_s^2\\
&- 24\,\mathrm{I}\left(1,0,0\right)m_bC_{Hgg}\,g_s^2 + 40s\,\mathrm{I}\left(1,0,1\right)m_bC_{Hgg}\,g_s^2 + 8\,\mathrm{I}\left(1,1,-1\right)m_bC_{Hgg}\,g_s^2\\
&- 16s\,\mathrm{I}\left(1,1,0\right)m_bC_{Hgg}\,g_s^2
\end{aligned} \tag{4.10}$$

This is a one-loop topology, so we know what the master integrals will look like: there can be triangle, bubble and tadpole integrals[1], and we will see this appear when reducing the family to a basis. For this

[1]Boxes are excluded because we only have three propagators

topology, we used FIRE5 to derive the IBP identities and reduce all the integrals in the amplitude to a set of three master integrals:



(4.11)

These integrals can be directly evaluated by integrating their Feynman parameter representations. They are of course well-known and can be found in a number of references, but there is a pedagogic interest in illustrating how they can be evaluated.

Direct integration of the master integrals

Evaluation of the tadpole integral The easiest integral is the tadpole integral $\mathrm{I}(0,0,1)$. We first need to evaluate its Feynman parametrization:

$$\mathrm{I}(0,0,1) = \int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2 - m_b^2} = \int_0^1 dx \delta\left(1-x\right) i (4\pi)^{-\frac{d}{2}} x^{1-d} \Gamma\left(1-\frac{d}{2}\right) \left(x^2 m_b^2\right)^{\frac{d}{2}-1} \tag{4.12}$$

No further work is required to find

$$\mathrm{I}(0,0,1) = i (4\pi)^{-\frac{d}{2}} \Gamma\left(1-\frac{d}{2}\right) \left(m_b^2\right)^{\frac{d}{2}-1} \tag{4.13}$$

Evaluation of the bubble integral The bubble integral $\mathrm{I}(1,1,0)$ requires slightly more work and in particular is a nice opportunity to showcase the role of the Cheng-Wu theorem in doing this kind of evaluation.

Let us first convert it to Feynman parametrization:

$$\mathrm{I}(1,1,0) = \int_0^\infty dx_1 dx_2 \delta\left(1-x_1-x_2\right) i (4\pi)^{-\frac{d}{2}} \left(x_1+x_2\right)^{2-d} \Gamma\left(2-\frac{d}{2}\right) \left(-s x_1 x_2\right)^{\frac{d}{2}-2} \tag{4.14}$$

This integrand is well-defined in the Euclidean region $s < 0$. We will always perform integrations in this region and then analyticaly continuate to the region $s > 0$ after integration. We can use the Cheng-Wu theorem to replace $\delta\left(1-x_1-x_2\right)$ by $\delta\left(1-x_2\right)$ and yield

$$\mathrm{I}(1,1,0) = \int_0^\infty dx_1 i (4\pi)^{-\frac{d}{2}} \left(1+x_1\right)^{2-d} \Gamma\left(2-\frac{d}{2}\right) \left(-s x_1\right)^{\frac{d}{2}-2}. \tag{4.15}$$

In which we can immediately identify an Euler beta function:

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = \int_0^\infty dz\, z^{a-1}(1+z)^{-a-b}, \tag{4.16}$$

and therefore

$$\mathrm{I}(1,1,0) = \frac{i(4\pi)^{-\frac{d}{2}}(-s)^{\frac{d}{2}-2}\Gamma\left(2-\frac{d}{2}\right)\Gamma\left(\frac{d}{2}-1\right)^2}{\Gamma(d-2)}. \tag{4.17}$$

Evaluation of the triangle integral The triangle integral $\mathrm{I}(1,1,1)$ requires using more technology as it will involve multiple polylogarithms. It is important that we clearly identify this integral for the upcoming two-loop calculation we will perform in the SM: we will see that in the $m_t \to \infty$ expansion, some two loop integrals appearing in the diagrams of Figure 4.4 will reduce to a form tadpole$\times$triangle.

The Feynman parametrization of this integral is

$$\mathrm{I}(1,1,1) = \int dx_1 dx_2 dx_3 \delta\left(1-\sum_{i\in\nu} x_i\right) i(4\pi)^{-\frac{d}{2}}\Gamma\left(3-\frac{d}{2}\right)(x_1+x_2+x_3)^{3-d}\left(x_3^2 m_b^2 - s x_1 x_2\right)^{\frac{d}{2}-3} \tag{4.18}$$

We can use the Cheng-Wu theorem to specify $\nu = \{3\}$ and define the form of the triangle integral that we will match in further calculations

$$\begin{aligned}
\mathrm{I}(1,1,1) =& \int_0^\infty dx_1 dx_2 i(4\pi)^{-\frac{d}{2}}\Gamma\left(3-\frac{d}{2}\right)(1+x_1+x_2)^{3-d}\left(m_b^2 - s x_1 x_2\right)^{\frac{d}{2}-3} \\
=& \int_0^\infty dx_1 dx_2 i(4\pi)^{-\frac{d}{2}}\Gamma\left(3-\frac{d}{2}\right) m_b^{d-6}(1+x_1+x_2)^{3-d}(1+r x_1 x_2)^{\frac{d}{2}-3} \\
=& i(4\pi)^{-\frac{d}{2}}\Gamma\left(3-\frac{d}{2}\right) m_b^{d-6}\,\mathrm{T}(r)
\end{aligned} \tag{4.19}$$

Where we defined $r = -s/m_b^2$ and the function

$$\mathrm{T}(r) = \int_0^\infty dx_1 dx_2 (1+x_1+x_2)^{3-d}(1+r x_1 x_2)^{\frac{d}{2}-3} \tag{4.20}$$

which we will now evaluate as a series in powers of our dimensional regularization parameter ϵ: $\mathrm{T}(r) = T_0(r) + \epsilon T_1(r) + \epsilon^2 T_2(r) + \mathcal{O}(\epsilon^3)$. We find

$$\begin{aligned}
T_0 =& \int_0^\infty dx_1 dx_2 \frac{1}{(x_1+x_2+1)(r x_1 x_2+1)} \\
T_1 =& \int_0^\infty dx_1 dx_2 \frac{2\log(x_1+x_2+1) - \log(r x_1 x_2+1)}{(x_1+x_2+1)(r x_1 x_2+1)} \\
T_2 =& \int_0^\infty dx_1 dx_2 \frac{\log^2(r x_1 x_2+1) - 4\log(x_1+x_2+1)\log(r x_1 x_2+1) + 4\log^2(x_1+x_2+1)}{2(x_1+x_2+1)(r x_1 x_2+1)}
\end{aligned} \tag{4.21}$$

The evaluation of these integrals can be handled in an algorithmic way, which we will illustrate on

T_0 and the first term of T_1. Since we expect the result to be expressed in terms of polylogarithms, the key element is to obtain integrals of the form $\int_0^a dx \frac{G(\ldots,x)}{x-b} = G(\ldots, b, a)$. This is done using repeated partial fractioning of the denominators. Let us start with the simplest case:

$$\begin{aligned} T_0 &= \int_0^\infty dx_1 dx_2 \frac{1}{(x_1+x_2+1)(rx_1x_2+1)} \\ &= \int_0^\infty dx_1 dx_2 \frac{rx_2}{(rx_1x_2+1)\left(rx_2^2+rx_2-1\right)} - \frac{1}{(x_1+x_2+1)\left(rx_2^2+rx_2-1\right)} \\ &= \int_0^\infty dx_2 \left[\frac{G(-x_2-1,x_1)}{-rx_2^2-rx_2+1} + \frac{G\left(-\frac{1}{rx_2},x_1\right)}{rx_2^2+rx_2-1}\right]_0^\infty \end{aligned} \tag{4.22}$$

Evaluating the primitive at the bound $x_1 = 0$ is easy in this case:

$$\lim_{x_1\to 0} \frac{G\left(-\frac{1}{rx_2},x_1\right) - G(-x_2-1,x_1)}{rx_2^2+rx_2-1} = 0, \tag{4.23}$$

Which results from the fact that $\lim_{x\to 0} G(\vec{a},x) = 0$ if $\vec{a} \neq \vec{0}$. In the other terms, we can encounter cases in which there is a logarithmic divergence at 0, *i.e.* polylogarithms of the form $G(\vec{0},x)$. These are handled using a series expansion instead of simply taking the limit, which then makes it obvious whether these logarithmic terms vanish in the limit ($\lim_{x\to 0} xG(\vec{0},x) = 0$) or are an actual divergence. Because the integrals are regularized and therefore overall convergent, all such divergences are generated by individual terms and cancel when combining them all. Evaluating the primitive at $x_1 = \infty$ is equivalent to taking the limit $\bar{x}_1 = 1/x_1 \to 0$. To obtain this limit, we use the algorithm described in Section 3.3.3 to reduce all polylogarithms of the form $G\left(\ldots,\frac{1}{\bar{x}_1}\right)$ into the canonical form $G(\ldots,\bar{x}_1)$, where taking the limit and extracting logarithmic divergences is straightforward. This yields the equalities:

$$\begin{aligned} G\left(-x_2-1,\tfrac{1}{\bar{x}_1}\right) &= -G(-1,x_2) - G(0,\bar{x}_1) + G\left(-\tfrac{1}{x_2+1},\bar{x}_1\right) \\ G\left(-\tfrac{1}{rx_2},\tfrac{1}{\bar{x}_1}\right) &= G(-rx_2,\bar{x}_1) + G(0,r) - G(0,\bar{x}_1) + G(0,x_2) \end{aligned} \tag{4.24}$$

Each polylogarithm yields a logarithmic divergence, which cancel against each other when combined into T_0. We therefore find

$$\mathrm{T}_0 = \int_0^\infty dx_2 \frac{G(0,r) + G(-1,x_2) + G(0,x_2)}{rx_2^2+rx_2-1} \tag{4.25}$$

Integrating this function into polylogarithms requires partial fractionning, which would generate unwieldy expressions with this choice of variables as the roots of the polynomial in the denominator of Equation 4.25 have a complicated relation to r:

$$rx_2^2+rx_2-1 = r\left(x_2 - \frac{-r-\sqrt{r+4}\sqrt{r}}{2r}\right)\left(x_2 - \frac{\sqrt{r}\sqrt{r+4}-r}{2r}\right). \tag{4.26}$$

We can however make a better choice by replacing r with $\mathfrak{r} = \frac{\sqrt{r}\sqrt{r+4}-r}{2r}$, the root of the denominator

that is positive in the Euclidean region. Indeed, the irreducible polynomials are then linear in $\mathfrak{r}$:

$$rx_2^2 + rx_2 - 1 = -\frac{(\mathfrak{r} - x_2)(\mathfrak{r} + x_2 + 1)}{\mathfrak{r}(\mathfrak{r}+1)}. \tag{4.27}$$

We can therefore rewrite T_0 in a form where the integration becomes straightforward:

$$\begin{aligned}\mathrm{T}_0 = \int_0^\infty dx_2 &\frac{\mathfrak{r}(\mathfrak{r}+1)\left(G\left(0,\frac{1}{\mathfrak{r}(\mathfrak{r}+1)}\right) + G(-1,x_2) + G(0,x_2)\right)}{(2\mathfrak{r}+1)(x_2-\mathfrak{r})} \\ &- \frac{\mathfrak{r}(\mathfrak{r}+1)\left(G\left(0,\frac{1}{\mathfrak{r}(\mathfrak{r}+1)}\right) + G(-1,x_2) + G(0,x_2)\right)}{(2\mathfrak{r}+1)(\mathfrak{r}+x_2+1)}\end{aligned} \tag{4.28}$$

Here the polylogarithms which depend on $1/(\mathfrak{r}(1+\mathfrak{r}))$ can be simplified easily because they are simple logarithms:

$$G\left(0,\frac{1}{r(r+1)}\right) \to -G(-1,r) - G(0,r), \tag{4.29}$$

but in general, it is a better choice to replace r by $\mathfrak{r}$ before applying the reduction to canonical basis and make it the last of the variables considered so that higher polylogarithms depending on $1/(\mathfrak{r}(1+\mathfrak{r}))$ be simplified directly, which is what we do when performing the complete integration. In any case, we reach the following form for T_0, for which we only need to repeat our trick to evaluate the primitive at infinity:

$$\begin{aligned}\mathrm{T}_0 = \Bigg[&\frac{\mathfrak{r}^2 G(-1,\mathfrak{r})G(-\mathfrak{r}-1,x_2)}{2\mathfrak{r}+1} + \frac{\mathfrak{r}^2 G(0,\mathfrak{r})G(-\mathfrak{r}-1,x_2)}{2\mathfrak{r}+1} - \frac{\mathfrak{r}^2 G(-1,\mathfrak{r})G(\mathfrak{r},x_2)}{2\mathfrak{r}+1} \\ &- \frac{\mathfrak{r}^2 G(0,\mathfrak{r})G(\mathfrak{r},x_2)}{2\mathfrak{r}+1} - \frac{\mathfrak{r}^2 G(-\mathfrak{r}-1,-1,x_2)}{2\mathfrak{r}+1} - \frac{\mathfrak{r}^2 G(-\mathfrak{r}-1,0,x_2)}{2\mathfrak{r}+1} + \frac{\mathfrak{r}^2 G(\mathfrak{r},-1,x_2)}{2\mathfrak{r}+1} \\ &+ \frac{\mathfrak{r}^2 G(\mathfrak{r},0,x_2)}{2\mathfrak{r}+1} + \frac{\mathfrak{r}G(-1,\mathfrak{r})G(-\mathfrak{r}-1,x_2)}{2\mathfrak{r}+1} + \frac{\mathfrak{r}G(0,\mathfrak{r})G(-\mathfrak{r}-1,x_2)}{2\mathfrak{r}+1} \\ &- \frac{\mathfrak{r}G(-1,\mathfrak{r})G(\mathfrak{r},x_2)}{2\mathfrak{r}+1} - \frac{\mathfrak{r}G(0,\mathfrak{r})G(\mathfrak{r},x_2)}{2\mathfrak{r}+1} - \frac{\mathfrak{r}G(-\mathfrak{r}-1,-1,x_2)}{2\mathfrak{r}+1} \\ &- \frac{\mathfrak{r}G(-\mathfrak{r}-1,0,x_2)}{2\mathfrak{r}+1} + \frac{\mathfrak{r}G(\mathfrak{r},-1,x_2)}{2\mathfrak{r}+1} + \frac{\mathfrak{r}G(\mathfrak{r},0,x_2)}{2\mathfrak{r}+1}\Bigg]_0^\infty\end{aligned} \tag{4.30}$$

At $x_2 = 0$, we can directly take the limit and find that the whole primitive evaluates to 0, since there are no logarithmic divergences. Replacing x_2 by $\bar{x}_2 = 1/x_2$ and reducing the polylogarithms to canonical form, we can easily take the limit $x_2 \to \infty$ and find:

$$\begin{aligned}\mathrm{T}_0 = &\frac{1}{3(2\mathfrak{r}+1)}\left(-3\mathfrak{r}^2 G(-1,\mathfrak{r})^2 + 3\mathfrak{r}^2 G(0,\mathfrak{r})^2 + 3\mathfrak{r}^2 G(-1,-1,\mathfrak{r})\right. \\ &- 3\mathfrak{r}^2 G(-1,0,\mathfrak{r}) + 3\mathfrak{r}^2 G(0,-1,\mathfrak{r}) - 3\mathfrak{r}^2 G(0,0,\mathfrak{r}) \\ &- 3\mathfrak{r}G(-1,\mathfrak{r})^2 + 3\mathfrak{r}G(0,\mathfrak{r})^2 + 3\mathfrak{r}G(-1,-1,\mathfrak{r}) \\ &\left.-3\mathfrak{r}G(-1,0,\mathfrak{r}) + 3\mathfrak{r}G(0,-1,\mathfrak{r}) - 3\mathfrak{r}G(0,0,\mathfrak{r}) + 2\pi^2\mathfrak{r}^2 + 2\pi^2\mathfrak{r}\right),\end{aligned} \tag{4.31}$$

which we can further simplify using the shuffle algebra:

$$\begin{aligned}\mathrm{T}_0 = &-\frac{\mathfrak{r}(\mathfrak{r}+1)G(-1,-1,\mathfrak{r})}{2\mathfrak{r}+1} - \frac{\mathfrak{r}(\mathfrak{r}+1)G(-1,0,\mathfrak{r})}{2\mathfrak{r}+1} + \frac{\mathfrak{r}(\mathfrak{r}+1)G(0,-1,\mathfrak{r})}{2\mathfrak{r}+1} \\ &+ \frac{\mathfrak{r}(\mathfrak{r}+1)G(0,0,\mathfrak{r})}{2\mathfrak{r}+1} + \frac{2\pi^2\mathfrak{r}(\mathfrak{r}+1)}{6\mathfrak{r}+3}\end{aligned} \tag{4.32}$$

Of course, following this procedure by hand for all terms in $\mathrm{T}(r)$ would be long and painstaking, but it is easy to streamline it once the algorithm for the reduction to the canonical form is implemented. For this, we used the private `Mathematica` package `PolyLogTools` [184]. The only other possible issue is that of having expressions linear in the mass scale ratio appearing in the polylogarithms so that the partial fractioning and the reduction algorithm can operate, but in our case, the replacement $r \to \mathfrak{r}$ is always sufficient. Here is the algorithm we applied to compute the triangle integral up to order $\mathcal{O}(\epsilon^2)$, which is essentially the method described in [159, 171]:

1. Start with T, an expression containing MPLs involving the variables $\{x_1, \dots, x_N\}$ and $\mathfrak{r}$. Set $j = 1$.
2. Reduce T to canonical form with respect to the variables $\{x_j, x_{j+1}, \dots, x_N, \mathfrak{r}\}$ (c.f. Section 3.3.3).
3. Partial fraction the coefficient of each polylogarithm in T with respect to x_j.
4. Find a primitive T_j of T with respect to x_j using $\displaystyle\int \frac{G(\vec{a}, x_j)}{b - x_j} = G(\vec{a}, b, x_j)$. The integral of T is $T_j^\infty - T_j^0$, the difference of the limits of T_j in ∞ and 0.
5. T_0 can be obtained as the series expansion of T in 0 up to $\mathcal{O}\left(x_j^0\right)$
6. To obtain T_∞, we will take the limit $1/x_j = \bar{x}_j \to 0$. Set $T_\infty = T(\bar{x}_j, x_{j+1}, \dots)$.
7. Reduce T_∞ to canonical form with respect to $\bar{x}_j, x_{j+1}, \dots, x_N, \mathfrak{r}$.
8. Expand T_∞ as a series up to $\mathcal{O}\left(\bar{x}_j^0\right)$.
9. Replace T by the integrated version $T_\infty - T_0$.
10. if $j = N$, end. Otherwise, set $j = j + 1$ and go to step 2.

Of course, this algorithm does not work systematically or we could evaluate any loop integral. The two crucial steps are step 2 and step 3: the set of entries of the polylogarithms and the rational prefactors in T need fullfill the criterium of linear reducibility discussed in Section 3.3.3 so that a canonical form can be reached after each variable is integrated.

Renormalized HEFT amplitude

Matching the HEFT to the SM can only be performed at the level of renormalized amplitudes. Indeed, the two-loop diagrams in Figure 4.4 are UV finite, while our one-loop triangle diagram $\mathcal{A}_{\text{gf}}$ is UV divergent:

$$\mathcal{A}_{\text{gf}} \ni -i\left(\frac{\alpha_s}{\pi}\right)^2 C_A C_F \left(4m_b^2 - s\right) \frac{m_b}{2v}\frac{1}{\epsilon}, \tag{4.33}$$

where we replaced C_{ggH} by its value and all the couplings are bare parameters. This is a manifestation of operator mixing: the Yukawa operator generates a counterterm for the gluon fusion operator.

In this subsection, we will keep using symbols without a superscript for bare couplings and will add a superscript R to renormalized parameters. We use the $\overline{\text{MS}}$ renormalization scheme for our calculation:

$$\begin{aligned}&\alpha_s = \mu^{2\epsilon} S_\epsilon^{-1} \alpha_s^R, \;\; v = \mu^{-\epsilon} v^R \\ &C_y = \mu^\epsilon \left(C_y^R + y_t^R \left(\frac{\alpha_s^R}{\pi}\right)^2 \left(\frac{m_b^R}{m_t^R}\right) \left(\frac{C_F}{2\epsilon} + \Delta_F\right)\right),\end{aligned} \tag{4.34}$$

where Δ_F is the finite quantity we wish to obtain using our matching and $S_\epsilon = (4\pi)^\epsilon e^{-\epsilon\gamma}$, with γ the Euler-Mascheroni constant. Note that the Higgs vacuum expectation value is not actually renormalized, but we must pay attention to separate the $(4-2\epsilon)$-dimensional quantity v to the 4-dimensional one $v^R = \mu^\epsilon v$ to ensure that each piece of the amplitude has the right dimension. Because we work with a single order in the perturbation theory, there is no need to specify the renormalization of the masses further than

$$m_t = m_t^R + \mathcal{O}\big(\alpha_s^R\big)\,,\; m_b = m_b^R + \mathcal{O}\big(\alpha_s^R\big)\,, \tag{4.35}$$

and we will suppress the subscript for the renormalized parameters used below.

We can now extract the piece of the HEFT amplitude for $H \to b\bar{b}$ of order $\mathcal{O}\big(y_t\alpha_s^2\big)$, which gets a contribution from $\mathcal{A}_b$ and from $\mathcal{A}_{\text{gf}}$:

$$\begin{aligned}&\mathcal{A}^{\text{HEFT}}_{H\to b\bar{b}}\big|_{y_t\alpha_s^2} = i y_t^R \left(\tfrac{\alpha_s^R}{\pi}\right)^2 C_A C_F \left(\frac{m_b}{m_t}\right) m_b^2 \frac{(2\mathfrak{r}+1)}{\mathfrak{r}(\mathfrak{r}+1)18\sqrt{2}} \Big(3G(0,-1,\mathfrak{r}) + 3G(0,0,\mathfrak{r}) \\ &- 3G(-1,-1,\mathfrak{r}) - 3G(-1,0,\mathfrak{r}) + 9(1+2\mathfrak{r})\log\left(\frac{m_b^2}{\mu^2}\right) - 24\mathfrak{r} + 2\pi^2 - 12\Big) \\ &+ i y_t^R \left(\frac{\alpha_s}{\pi}\right)^2 \left(\frac{m_b}{m_t}\right) C_A \frac{m_b^2}{\sqrt{2}} \frac{(2\mathfrak{r}+1)^2}{\mathfrak{r}(\mathfrak{r}+1)} \Delta_F.\end{aligned} \tag{4.36}$$

4.2 Higgs-to-bottom decay at two loops in the Standard Model

Let us now move on to the Standard Model calculation of $H \to b\bar{b}$ at $\mathcal{O}\big(\alpha_s^2 y_t\big)$. The two diagrams entering the calculation are shown in Fig. 4.4. As in the HEFT, we are computing their spin and color summed contraction with spinors to deal only with scalar quantities. We will proceed through the same steps as in the previous section, defining a topology and reducing the integral to master integrals using IBP relations, before evaluating the master integrals. The major difference with the previous section is that we now need to expand this amplitude in the large m_t limit. To do so, we will reduce the exact amplitude to master integrals, compute the expansion of the masters in the large m_t limit using the method of regions and expand their coefficients in the amplitude.

4.2.1 Reducing the amplitude to master integrals

For the evaluation of the diagrams in this section, we used `QGRAF` to generate the diagrams in my `XML` format and then translated them into `FORM` for evaluation of the fermion traces. We then imported these results in `Mathematica`, in which we performed the rest of the calculations.

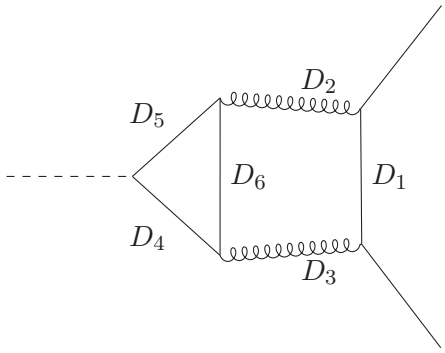


Figure 4.7: Correspondence between the denominators for our topology and the denominator in one of our Feynman diagrams.

As in the case of HEFT, the presence of only two diagrams made the choice of a topology straightforward. We opted for the following:

$$\begin{aligned} &D_1 = k_2^2 - m_b^2, && D_2 = (k_2 - p_1)^2, && D_3 = (k_2 + p_2)^2, \\ &D_4 = (k_1 - p_1)^2 - m_t^2, && D_5 = (k_1 + p_2)^2 - m_t^2, && D_6 = (k_1 - k_2)^2 - m_t^2, \\ & && D_7 = (k_1 + k_2)^2 - m_t^2. \end{aligned} \tag{4.37}$$

Their relation to the Feynman diagrams is shown in Figure 4.7. Note that an extra propagator D_7 needs to be added to the definition of the topology so as to deal with all possible numerators in integrals: there are 7 possible scalar products involving k_1 or k_2 so we need 7 propagators to deal with them all.

Using these definitions, we can derive the rules for the scalar products replacements:

$$\begin{aligned} &k_1 \cdot k_1 = -D_1 + \tfrac{1}{2}D_6 + \tfrac{1}{2}D_7 + m_t^2 - m_b^2, && k_2 \cdot k_2 = D_1 + m_b^2, \\ &k_1 \cdot p_1 = -\tfrac{1}{2}D_4 + \tfrac{1}{4}D_6 + \tfrac{1}{4}D_7 - \tfrac{1}{2}D_1, && k_2 \cdot p_1 = \tfrac{1}{2}D_1 - \tfrac{1}{2}D_2 + m_b^2, \\ &k_1 \cdot p_2 = \tfrac{1}{2}D_5 - \tfrac{1}{4}D_6 - \tfrac{1}{4}D_7 + \tfrac{1}{2}D_1, && k_2 \cdot p_2 = -\tfrac{1}{2}D_1 + \tfrac{1}{2}D_3 - m_b^2, \\ &k_1 \cdot k_2 = \tfrac{1}{4}D_7 - \tfrac{1}{4}D_6. \end{aligned} \tag{4.38}$$

We can now define any integral in our amplitude as a linear combinations of integrals of the form:

$$J(n_1, \ldots, n_7) = \int \frac{d^d k_1}{(2\pi)^d} \frac{d^d k_2}{(2\pi)^d} \frac{1}{D_1^{n_1} \ldots D_7^{n_7}}. \tag{4.39}$$

As in the HEFT calculation, we obtained the expresssion of the two diagrams in terms of topology integrals going through the chain of `QGRAF`, `qgraf-xml-parser`, `FORM` and `SAIF`. We used `FIRE5` to

reduce the 28 integrals appearing in these diagrams to 23 master integrals which are listed below:

$$\begin{array}{llll} J(0,0,0,1,0,1,0), & J(1,0,0,1,0,0,0), & J(0,0,1,1,0,1,0), & J(0,0,2,1,0,1,0), \\ J(0,0,1,2,0,1,0), & J(0,0,0,1,1,1,0), & J(0,1,1,1,0,0,0), & J(2,0,0,0,1,1,0), \\ J(1,0,0,1,0,1,0), & J(2,0,0,1,0,1,0), & J(1,0,0,1,1,0,0), & J(0,1,1,1,1,0,0), \\ J(1,0,0,1,1,1,0), & J(2,0,0,1,1,1,0), & J(1,0,0,1,1,2,0), & J(1,0,0,2,1,1,0), \\ J(1,0,1,1,0,1,0), & J(2,0,1,1,0,1,0), & J(1,0,1,2,0,1,0), & J(1,1,1,1,0,0,0), \\ J(0,1,1,1,1,1,0), & J(1,1,1,1,1,0,0), & J(1,1,1,1,1,1,0). & \end{array} \tag{4.40}$$

4.2.2 Evaluation of the master integrals in the large top mass limit

Our goal is to compute the expansion of the SM amplitude in the limit $m_t \to \infty$, so we do not need to evaluate these integrals exactly. Instead, we expand them in this limit using the method of regions and integrate the expanded integrals, which simplifies the problem greatly. To this end, we used the `Mathematica` package `UF` to obtain the Feynman parameter expression for each of these integrals and then used the package `ASY` to find relevant regions and expand the integrals in each relevant region.

Technically speaking, the expansion is set up by introducing a spurious parameter ρ in which we will expand integrals and assigning each external parameter a scaling in ρ. Positive powers of ρ are assigned to parameters that are suppressed in the limit of interest so we will assign positive powers to m_b and s and a trivial scaling to m_t. We can always shift the powers of ρ homogeneously, so this choice will produce exactly the same result as if we had sent $m_t \to \infty$ and kept m_b and s fixed, but it has the practical advantage of yielding a series in ρ which starts at a non-negative power. The scaling we chose is the following:

$$m_t \to \rho^0 m_t, \quad m_b \to \rho m_b, \quad s \to \rho^2 s. \tag{4.41}$$

Each region is then defined by the scaling of each Feynman parameter in ρ: $x_i \to \rho^{\kappa_i} x_i$. The relative scaling of the parameters determines whether the corresponding propagator in the integral is hard or soft. As discussed in Section 3.2.2, this provides a nice diagrammatic representation of the integral in each region, which will help us understand how the matching works.

The package `ASY` provides the list of regions that contribute to each integral in the limit we are considering. For each region, we rescale external and Feynman parameters accordingly and then expand the rational function in the integrand in powers of ρ without modifying the delta function, which is rigorous as we discussed in Section 3.2.2.

We have expanded all 23 master integrals and integrated them directly using the methods discussed in the previous section. Their expressions, which we have checked numerically against `FIESTA`, are shown Appendix C, but it is interesting to see on an example how the expansion and integration work. The techniques used for integrations are the same as the ones presented in Section 4.1, augmented by another trick: it is often useful to apply changes of variables of the form $x_i \to x_i x_j$ to make gamma functions apparent. For example, applying $x_1 \to x_1 x_2$,

$$\int_0^\infty dx_1 dx_2\, x_1 (x_1+x_2)^\alpha (1+x_2)^\beta = \int_0^\infty dx_1 dx_2\, x_1 (1+x_1)^\alpha x_2^{\alpha+2} (1+x_2)^\beta \\ = \frac{\Gamma(-\alpha-2)\Gamma(\alpha+3)\Gamma(-\alpha-\beta-3)}{\Gamma(-\alpha)\Gamma(-\beta)}. \tag{4.42}$$

Let us consider the most complicated master integral in our basis: $J(1,1,1,1,1,1,0)$, which can be diagrammatically represented as in Figure 4.7. Having six propagators and three scales, the Feynman parameter representation of this integral is quite involved:

$$\begin{aligned} &J(1,1,1,1,1,1,0) = -\pi^d \Gamma(6-d) \\ &\int \prod dx_i \left(x_1x_4 + x_2x_4 + x_3x_4 + x_6x_4 + x_1x_5 + x_2x_5 + x_3x_5 + x_1x_6 + x_2x_6 + x_3x_6 + x_5x_6\right)^{6-\frac{3d}{2}} \\ &\times \left(x_1^2x_4m_b^2 + x_1^2x_5m_b^2 + x_1^2x_6m_b^2 + x_1x_4^2m_t^2 + x_2x_4^2m_t^2 + x_3x_4^2m_t^2 + x_1x_5^2m_t^2 + x_2x_5^2m_t^2 + x_3x_5^2m_t^2 \right. \\ &+x_1x_6^2m_t^2 + x_2x_6^2m_t^2 + x_3x_6^2m_t^2 + x_4x_6^2m_t^2 + x_5x_6^2m_t^2 + 2x_1x_4x_5m_t^2 + 2x_2x_4x_5m_t^2 + 2x_3x_4x_5m_t^2 \\ &+x_4^2x_6m_t^2 + x_5^2x_6m_t^2 + 2x_1x_4x_6m_t^2 + 2x_2x_4x_6m_t^2 + 2x_3x_4x_6m_t^2 + 2x_1x_5x_6m_t^2 + 2x_2x_5x_6m_t^2 \\ &+2x_3x_5x_6m_t^2 + 2x_4x_5x_6m_t^2 - sx_2x_3x_4 - sx_2x_3x_5 - sx_1x_4x_5 - sx_2x_4x_5 - sx_3x_4x_5 \\ &\left. -sx_2x_3x_6 - sx_3x_4x_6 - sx_2x_5x_6 - sx_4x_5x_6\right)^{d-6}, \end{aligned} \tag{4.43}$$

where x_i is the Feynman parameter associated to denominator D_i. Processing this integral through ASY, we find that two regions contribute to our expansion.

HARD REGION In the first region, the scaling of the Feynman parameters is uniform: $x_i \to \rho^0 x_i$. This corresponds to setting all loop momenta hard. In the diagrammatic picture described in Section 3.2.2, the complete two loop diagram for the integral (Figure 4.7) is reduced to a single point. The functional dependence of the integral in this region highlights this fact: the only parameters proportional to some power of ρ are s and m_b. As a result, the contribution from this region takes the form

$$J(1,1,1,1,1,1,0)\big|_{\text{hard}} = \sum_{n=0}^{\infty} m_b^n \sum_{k=k_{\min}^{(n)}}^{k_{\max}^{(n)}} \left(\frac{s}{m_b^2}\right)^k \int \prod_{i=1}^{6} dx_i J_h^{(nk)}(x_i; m_t) \tag{4.44}$$

Each integrand in the expansion only depends on the top mass, which means that they essentially are tadpole integrals and they correspond to localized interactions, confirming the picture that this region generates the part of the amplitude which will be matched to the tree-level part of our HEFT amplitude. That these integrals are tadpoles also shows in the simplicity of the integration, which can be performed readily with our list of tricks. Let us derive the first order in the expansion, $J_h^{(00)}$.

$$\begin{aligned} J_h^{(00)} &= -\pi^d \Gamma(6-d) m_t^{2(d-6)} \delta\left(\sum_{i\in\nu} x_i - 1\right) (x_4+x_5+x_6)^{d-6} \\ &\left(x_3x_4 + x_6x_4 + x_3x_5 + x_3x_6 + x_5x_6 + x_1(x_4+x_5+x_6) + x_2(x_4+x_5+x_6)\right)^{-d/2}. \end{aligned} \tag{4.45}$$

We can use the Cheng-Wu theorem to specify $\nu = \{1\}$ and apply the rescalings $x_4 \to x_4x_6$, $x_5 \to$

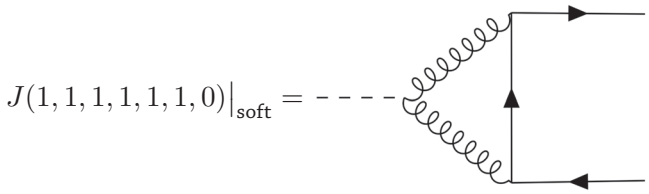


Figure 4.8: Diagrammatic interpretation of the soft contribution to the integral $J(1,1,1,1,1,1,0)$

$x_5 x_6$ to factorize x_6 out of the first term. The integral over x_6 takes the form $\int dx_6\, x_6^{a_1}(b_1+b_2x_6)^{a_2}$, which is a Beta function integral. At this stage, we can directly integrate x_2 and x_3:

$$\begin{aligned}\int\prod dx_i J_h^{(00)} &= -16\pi^d m_t^{2(d-6)}\Gamma(6-d)\int\prod dx_i \frac{(x_4+x_5)^{3-\frac{d}{2}}\,(x_4+x_5+1)^{d-9}}{(d-6)(d-4)(d-2)\,(x_2+x_3+1)^3}\\ &= -8\pi^d m_t^{2(d-6)}\Gamma(6-d)\int\prod dx_i \frac{(x_4+x_5)^{3-\frac{d}{2}}\,(x_4+x_5+1)^{d-9}}{(d-6)(d-4)(d-2)}\end{aligned} \tag{4.46}$$

We are now in the same situation as in the first step: by factorizing x_5 out of the first term through the rescaling $x_4 \to x_4 x_5$, we make the integral over x_5 take the form of a Beta function and leaves the integral over x_4 trivial:

$$\begin{aligned}\int\prod dx_i J_h^{(00)} &= -8\pi^d m_t^{2(d-6)}\Gamma(6-d)\frac{(x_4+1)^{3-\frac{d}{2}}\,x_5^{4-\frac{d}{2}}\,((x_4+1)\,x_5+1)^{d-9}}{(d-6)(d-4)(d-2)}\\ &= 2^{d-8}\pi^{d+\frac{1}{2}} m_t^{2(d-6)}\frac{\Gamma(6-d)\Gamma\left(1-\frac{d}{2}\right)}{\Gamma\left(\frac{9}{2}-\frac{d}{2}\right)}\int dx_4 \frac{1}{(x_4+1)^2}\\ &= \frac{2^{d-8}\pi^{d+\frac{1}{2}}\Gamma(6-d)\Gamma\left(1-\frac{d}{2}\right)m_t^{2(d-6)}}{\Gamma\left(\frac{9}{2}-\frac{d}{2}\right)}\end{aligned} \tag{4.47}$$

Soft region The second region has non uniform scaling in the Feynman parameters:

$$\begin{array}{l|l} x_1 \to \rho^2 x_1 & x_4 \to \rho^0 x_4 \\ x_2 \to \rho^2 x_2 & x_5 \to \rho^0 x_5 \\ x_3 \to \rho^2 x_3 & x_6 \to \rho^0 x_6 \end{array} \tag{4.48}$$

The diagrammatic interpretation of this scaling is that $x_1 \dots x_3$ are soft and therefore unpinched while $x_4 \dots x_6$ are hard and therefore contracted to a point. As a result, we can make a representation of the soft region contribution $J(1,1,1,1,1,1,0)\big|_{\text{soft}}$ of the integral as shown in Figure 4.8.

This description clearly suggests that this soft region part of our integral should contribute to the part of the HEFT amplitude corresponding to the one-loop diagram in Figure 4.5. We will see that it is indeed the case as it will actually be proportional to the triangle integral we have computed in Section 4.1.

After introducing the scaling in ρ of the parameters in this region in the expression of $J(1,1,1,1,1,1,0)$, we can expand the integral, which now has a nontrivial dependence on the kinematics of the process:

$$J(1,1,1,1,1,1,0)\big|_{\text{soft}} = \sum_{n=n_{\min}}^{\infty} m_b^n \int \prod_i^6 J_s^{(n)}(x_i; m_t, s/m_b^2), \tag{4.49}$$

which will eventually allow the integrals $J_s^{(n)}$ to see threshold effects when s crosses $4m_b^2$. The first nontrivial term is the integral $J_s^{(-2)}$. This term diverges in the limit $m_b \to 0$ so its contribution will actually cancel against other terms in the integral. This means that we will need to expand this integral to higher orders in the expansion, which we actually need to do for several integrals. For the purpose of illustrating the evaluation of integrals in our calculation, it is however still useful to go through the steps of the derivation.

$$\begin{aligned}\int \prod dx_i J_s^{(-2)} &= \pi^d m_t^{2(d-7)} \Gamma(7-d) \int \prod dx_i \delta\left(\sum_{i\in\nu} x_i - 1\right) (x_4+x_5+x_6)^{d-7} \\ &\quad ((x_2+x_3+1)(x_4+x_5) + (x_2+x_3+x_4+x_5+1)\,x_6)^{-\frac{d}{2}-1} \left[m_b^2 (x_4+x_5+x_6)\right. \\ &\quad \left. -s\,(x_4\,(x_3x_6 + x_5\,(x_3+x_6+1)) + x_2\,(x_5\,(x_4+x_6) + x_3\,(x_4+x_5+x_6)))\right]\end{aligned} \tag{4.50}$$

The integration is performed for all parameters save x_2, x_3 using the usual tricks and we find

$$\begin{aligned}\int \prod dx_i J_s^{(-2)} &= -\tfrac{1}{2}\pi^d (m_b m_t)^{d-6} \Gamma\left(3-\tfrac{d}{2}\right)^2 \int dx_2 dx_3\ (1+x_2+x_3)^{3-d} (1+r x_2 x_3)^{\frac{d}{2}-3} \\ &= -\tfrac{1}{2}\pi^d (m_b m_t)^{d-6} \Gamma\left(3-\tfrac{d}{2}\right)^2 \mathrm{T}(r),\end{aligned} \tag{4.51}$$

in which the triangle integral we computed in Section 4.1 appears naturally, confirming the intuition that this is a contribution to the part of the amplitude that will be matched to the one-loop diagram in the HEFT.

Second order correction to the expansion The leading order contribution from some integral yields a divergent contribution to the form factor in the limit $m_t \to \infty$, all of which cancel against each other when combined. As a result, several integrals, including $J(1,1,1,1,1,1,0)$, need to be expanded beyond the leading order in the $m_t \to \infty$ limit in order to obtain their contribution to the first non-vanishing order of the expansion of the form factor. Obtaining the integrand for these corrections is no more complicated as in the leading order: it amounts to taking the second nontrivial term in the series expansion of the integrand in powers of ρ. All the integrals thus obtained can be easily integrated using the set of tricks layed out in the previous discussion except for the NLO correction to the soft region of $J(1,1,1,1,1,1,0)$, $J_s^{(0)}$. As in the leading-order case, all parameters can be integrated

straightforwardly except x_2 and x_3, and the remaining integral is the following:

$$\begin{aligned} J_s^{(0)} &= -\tfrac{1}{48}\pi^d m_b^{d-4} m_t^{d-8} \Gamma\left(2-\tfrac{d}{2}\right)\Gamma\left(4-\tfrac{d}{2}\right)\int dx_1 dx_2\ (x_2+x_3+1)^{-d-1}\,(r x_2 x_3+1)^{\frac{d}{2}-3} \\ &\quad \left[d\left(r\left(x_2\left(2x_2+3\right)+\left(x_3+1\right)\left(2x_3+1\right)\right)-4\right)\right. \\ &\quad \left.-4r\left(2x_2^2+\left(2x_3+3\right)x_2+x_3\left(2x_3+3\right)\right)-4r+8\right] \end{aligned} \tag{4.52}$$

While this integral is more complicated than $\mathrm{T}(r)$, it can be expanded in powers of $\epsilon = (4-d)/2$ and evaluated using the same algorithm. The result of the integration can be found in Appendix C.

4.2.3 The expanded Standard Model amplitude

We now have all the elements necessary for the evaluation of the Feynman diagrams for $H \to b\bar{b}$ at $\mathcal{O}\left(\alpha_s^2 y_t\left(m_b/m_t\right)\right)$ in the Standard Model. Since we have performed the reduction to master integrals in the exact amplitude, inserting their expanded values yields an expression of the form

$$\mathcal{A}_{\text{SM}} = \sum_i A_i(m_t, m_b, s)\left(\sum_{n=n^i_{\min}}^{n^i_{\max}}\left(\frac{m_b}{m_t}\right)^n \sum_{k=k_n^{\min,i}}^{k_n^{\max,i}}\left(\frac{s}{m_b^2}\right)^k f_i^{nk}(m_t, m_b, s)\right), \tag{4.53}$$

where the A_i are rational functions of the external scales, the f_i^{nk} have no power-law dependence on m_b and s, and i spans each region of each integral. To obtain $\mathcal{A}_{\text{SM}}$ as a power series in m_b and s, we now only need to expand the A_i to a sufficient order. We can check for missing orders in the expansion of the master integrals by checking that an extra order in the expansion of the integral would not modify the expanded form factor $\mathcal{A}_{\text{SM}}$. In practice, we add a dummy variable to each master integral, representing the neglected power suppressed term, and check that these dummy variables vanish when we expand the form factor. After performing this task using the first order of the expansion for all the master integrals, we found that we needed the power suppressed terms of some master integrals and we therefore computed them.

While the amplitude in the SM is finite, it is still relevant to specify a renormalization scheme so that unwanted factors of $\log 4\pi$ and the Euler-Mascheroni γ constant be removed, and the dimensionality of the amplitude be respected. In our case we only need specify

$$\alpha_s = \mu^{2\epsilon} S_\epsilon^{-1} \alpha_s^R, \quad v = \mu^{-\epsilon} v^R. \tag{4.54}$$

As in the HEFT, the renormalization of the masses yields a trivial change so we avoid using superscripts below.

The final result for the SM amplitude for $H \to b\bar{b}$ at $\mathcal{O}\left(y_t^R \alpha_s^{R2}\left(m_b/m_t\right)\right)$ in terms of renormalized parameters is

$$\begin{aligned} \mathcal{A}_{\text{SM}} &= i y_t^R \left(\alpha_s^R\right)^2 C_A C_F \left(\frac{m_b}{m_t}\right)\frac{m_b^2}{36\sqrt{2}\pi^2}\frac{(2\mathfrak{r}+1)}{\mathfrak{r}(\mathfrak{r}+1)}\left(6G(0,-1,\mathfrak{r})+6G(0,0,\mathfrak{r})\right. \\ &\quad \left.-6G(-1,-1,\mathfrak{r})-6G(-1,0,\mathfrak{r})+36\left(2\mathfrak{r}+1\right)\log\left(\frac{m_b}{m_t}\right)-78\mathfrak{r}+4\pi^2-39\right). \end{aligned} \tag{4.55}$$

4.3 Extraction of the m_b/m_t correction to the bottom Yukawa

We now have all the necessary elements to extract the $\mathcal{O}(m_b/m_t)$ correction to the HEFT bottom Yukawa coupling by matching the contributions of order $\mathcal{O}\big(y_t^R \alpha_s^{R2}(m_b/m_t)\big)$ to the $H \to b\bar{b}$ amplitude in the SM and in the HEFT. By imposing that the two amplitudes be equal up to corrections suppressed by powers of the top mass, we can fix the value of Δ_F. Indeed, the difference between the two contributions we have calculated is

$$iy_t^R \left(\frac{\alpha_s^R}{\pi}\right)^2 \left(\frac{m_b}{m_t}\right) C_A m_b^2 \frac{(2\mathfrak{r}+1)^2}{\mathfrak{r}(\mathfrak{r}+1)} \left(\frac{6C_F \log\left(\frac{m_t^2}{\mu^2}\right) + 5C_F + 12\Delta_F}{12\sqrt{2}}\right). \tag{4.56}$$

All the parameters in this equations are renormalized parameters evaluated at the renormalization scale μ. The matching condition for Δ_F is therefore:

$$\Delta_F(\mu) = -\frac{C_F}{12}\left(5 + 6\log\left(\frac{m_t^2}{\mu^2}\right)\right), \tag{4.57}$$

which we can translate to the value of the bottom Yukawa coupling in the HEFT at NLO in the $1/m_t$ expansion:

$$C_y(\mu) = y_b^R - y_t^R \left(\frac{\alpha_s^R}{\pi}\right)^2 \left(\frac{m_b}{m_t}\right) \frac{C_F}{12}\left(5 + 6\log\left(\frac{m_t^2}{\mu^2}\right)\right). \tag{4.58}$$

5

Gluon fusion in the SMEFT at two loops

The experimental collaborations at the LHC have delivered a comprehensive account of the properties of the boson discovered in 2012, showing that it behaves as expected from a Standard Model Higgs boson within uncertainties: it has the right spin and parity [80, 81], as well as the right interactions [7]. The current experimental precision leaves room for substantial deviations in the couplings of the Higgs boson to other Standard Model particle and it is our job as theorists to provide the tools that will enable us to extract either the most useful information from a signal, or the most meaningful constraints if we find no new physics.

The information extracted from the Run I on Higgs boson couplings by ATLAS and CMS is formulated in the κ-framework [45], which we discussed in Section 2.3.1. The κ parameters are fit parameters which can be intuitively interpreted as modified couplings and their definitions are based on the leading-order parametric dependence of Higgs production and decay on Higgs couplings. While this is an intuitive approach to putting constraints on possible new physics effects on Higgs couplings, the fact that the κ are not field-theoretic parameters makes their link to actual BSM models complicated. Extracting a precise relation between the κ and the parameters of a BSM model essentially requires calculating all the processes considered in the experimental signal fits to a sufficient degree of accuracy. This involves making many calculations in each BSM model at least at NLO, as experience has shown that LO predictions fall far from the mark in Higgs physics.

An alternative approach to describe possible new physics effects in a model-independent way is to use the SMEFT [60–62]. This makes the relation between the signal and the model-independent description more complicated, but the relation of BSM parameters to SMEFT parameters is much simpler than to κ parameters. This is a desirable trade-off as the SMEFT calculation needs only be done once: we can expand the effort to provide precise SMEFT predictions for key processes and extract experimental constraints on SMEFT parameters which can then be translated into constraints on BSM models.

A number of LHC processes has been studied in the SMEFT already, in particular in top physics, where NLO QCD predictions are available for $t\bar{t}$ [185], single top [117] and $t\bar{t}H$ [118] productions, as well as decays [120]. Precise predictions also exists in Higgs physics, for decays [121, 122, 186] and associated production [124], but so far only parts of the gluon fusion cross-section are known at NLO [123, 187]. As we have already stressed, NLO corrections in Higgs physics are particularly important [51, 52] and the LO prediction in the SMEFT is potentially unreliable. We therefore undertook the calculation of the NLO corrections to Higgs boson production in gluon fusion in the SMEFT.

Gluon fusion is a loop-induced process in the SM, dominated by loops of top quarks. We set up our calculation in order to be sensitive not only to the effects of heavy particles coupling gluons and Higgs bosons, but also to new interactions involving top quarks, Higgs bosons and gluons which could modify the top loop. As a result, the LO process in the SMEFT is also loop-induced and the virtual corrections to this process are two-loop diagrams. This chapter will describe how we performed the calculation of these virtual corrections using modern multi-loop techniques and the next will discuss the phenomenology of the gluon fusion process at NLO. The discussion of the calculation is organized as follows: Section 5.1 details how we set up the calculation in the SMEFT and discusses the relevant operators, Section 5.2.3 explains how we evaluated the two-loop diagrams, Section 5.3 discusses the divergence structure of the amplitude and presents the final, renormalized and IR-subtracted result. We then derive the renormalization group evolution of the Wilson coefficients in Section 5.4 and perform the analytic continuation to the physical regions in Section 5.5.

5.1 The gluon fusion process in the SMEFT

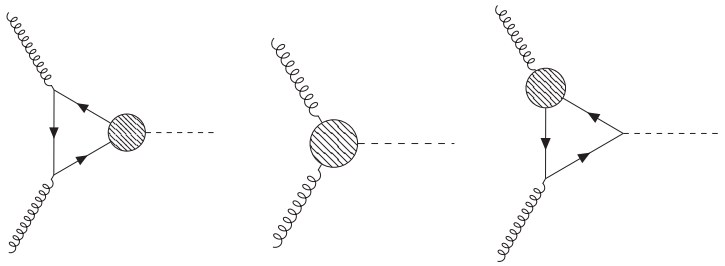
The SMEFT is the extension of the Standard Model by the tower of higher dimensional operators involving SM fields:

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_{i=1}^{\infty} \sum_{j=0}^{N_i} \frac{c_j^{(i)}}{\Lambda^i} O_j^{(i)}, \tag{5.1}$$

where the operators $O_j^{(i)}$ have an energy dimension of $4+i$, Λ is an energy scale larger than typical LHC energies and $c_j^{(i)}$ are dimensionless couplings. While non-renormalizable in general, the SMEFT Lagrangian can be safely truncated at some order $1/\Lambda^n$ and the resulting set of operators is stable under renormalization as long as observables are also truncated at that order [86]. For most practical purposes, it is sufficient to work at the order $1/\Lambda^2$: it is the first non-trivial order which features operators that affect virtually all aspects of LHC physics. A dimension 5 operator exists but is related to neutrino physics [60] and bears little effect on collider observables. Since we will work at a fixed order in the expansion in $1/\Lambda$, we will absorb it in the definition of our Wilson coefficient and work with the dimensionful coupling

$$C_i = \frac{c_i}{\Lambda^2}. \tag{5.2}$$

In this chapter, we will always discuss the case of operators involving the top quark. It should however be clear that in principle we should consider the contribution of operators involving up- and down-type quarks of all three generations as our goal is to provide a result to be used in global fits of the SMEFT and we cannot arbitrarily choose to set any operator to 0. Nevertheless, the analytic calculation of the



(a) O_1 **(b)** O_2 **(c)** O_3

Figure 5.1: Sample diagrams in the leading order contribution to gluon fusion for O_1, O_2, and O_3. The blobs indicate the dimension-6 operators.

two-loop amplitude is independent of the nature of the quark running in the loop so we will take it to be the top quark by default throughout this chapter and leave the discussion of the physical relevance of each quark flavor to Chapter 6, in which we discuss the phenomenology of the process. The only section where flavor plays a role is in the analytic continuation of the amplitude from the Euclidean region to either of the two possible physical region: below the quark pair-production threshold (for the top quark) and above the threshold (for the light flavors), which will be discussed in Section 6.4.

Three CP-conserving, dimension-six operators affect gluon fusion by modifying the top loop [187]:

$$O_1 = \bar{Q}_L \phi t_R (\phi^\dagger \phi - \frac{v^2}{2}) + \text{h.c.}\,, \tag{5.3}$$

$$O_2 = g_s^2 G^a_{\mu\nu} G_a^{\mu\nu} (\phi^\dagger \phi - \frac{v^2}{2}) + \text{h.c.}\,, \tag{5.4}$$

$$O_3 = g_s \bar{Q}_L T_a \sigma^{\mu\nu} t_R G^a_{\mu\nu} \phi + \text{h.c.}\,, \tag{5.5}$$

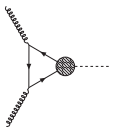
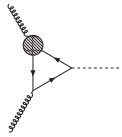
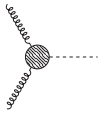
where $\sigma^{\mu\nu} = i\left[\gamma^\mu, \gamma^\nu\right]/\sqrt{2}$. Their leading-order contribution to this process is exemplified in Figure 5.1. These operators are normalized so that their first contribution to gluon fusion appears at the same order in α_s as the SM LO. We subtract vacuum expectation values from the operators to avoid redefining SM fields and couplings and we add the Hermitian conjugate of all operators, even to O_2 which is already Hermitian, to match the conventions of [118].

The contribution of these operators has been studied at LO in [187], and the NLO+NLL contribution from O_1 and O_2 was computed in [123]. The LO result underlined the complexity associated with O_3, which generates an intricate UV divergence structure. Indeed, already at leading order, the diagrams generated by O_3, exemplified in Figure 5.1c, are UV divergent and requires a counterterm in C_2:

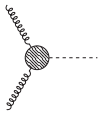
$$C_2^B \ni C_3^R \frac{\delta Z_{23}}{\epsilon}, \tag{5.6}$$

as illustrated schematically in Equation 5.7 which describes the structure of the part of the amplitude generated by the EFT operators M_{EFT}. This is the a manifestation of operator mixing which also ap-

peared in the previous chapter.

$M_{\text{EFT}} =$ $C_1 \times$  $+C_3 \times$  $+C_2 \times$ 

Finite UV divergent Finite

$+\delta \frac{Z_{23}}{\epsilon} C_3 \times$ 

Counter-term (5.7)

5.2 Calculation of the amplitude

5.2.1 Structure of the amplitude

The $gg \to H$ process has a very simple tensor and color structure that makes it easy to go from the full amplitude to a scalar quantity that we can efficiently deal with. The amplitude takes the form

$$\begin{aligned} \mathcal{A}(p_1, \sigma_1, c_1, p_2, \sigma_2, c_2) &= \varepsilon_\mu(p_1, \sigma_1)\varepsilon_\nu(p_1, \sigma_1)\mathcal{A}^{\mu\nu}_{\text{amp}\, c_1 c_2} \\ &= \varepsilon_\mu(p_1, \sigma_1)\varepsilon_\nu(p_1, \sigma_1)\delta_{c_1 c_2} T^{\mu\nu}\mathcal{M}, \end{aligned} \tag{5.8}$$

where the ε are the polarization vectors of the gluons, the p_i are their momenta, σ_i are their helicities and c_i are their colors. The tensor structure is $T^{\mu\nu} = (p_1 \cdot p_2 g^{\mu\nu} - p_1^\nu p_2^\mu)$. We will use the scalar invariant $s = (p_1 + p_2)^2 = 2p_1 \cdot p_2$ throughout this calculation to cleanly disentangle the physical Higgs boson mass $m_H = 125\ \text{GeV}$ and the variable s we use in our calculation. Until we discuss the analytic continution in Section 5.5, we will always assume to be in the Euclidean region ($s < 0$, $m_t^2 > 0$). We regularize the amplitudes with dimensional regularization and use the regularization parameter $\epsilon = (4 - d)/2$ where d is the number of spacetime dimensions.

We can obtain the scalar amplitude $\mathcal{M}$ by applying a projector $\mathcal{P}^{c_1 c_2}_{\mu\nu}$ to the amplitude amputated of its external states $\mathcal{A}_{\text{amp}}$. The required projector is

$$\mathcal{P}^{c_1 c_2}_{\mu\nu} = \frac{2}{s^2(1-\epsilon)} T^{\mu\nu} \frac{\delta^{c_1 c_2}}{N_c^2 - 1}. \tag{5.9}$$

We will decompose the bare amplitude $\mathcal{M}^B$ as follows:

$$\begin{aligned} \mathcal{M}^B(s, m_t^B) = \frac{\alpha_s^B}{\pi}\mu^{-2\epsilon} S_\epsilon \left(\frac{1}{v^B} M^B_{\text{SM}} + C_1^B \left(v^B\right)^2 M_1^B + C_2^B v^B M_2^B + C_3^B M_3^B \right) \\ + \left(\frac{\alpha_s^B}{\pi}\mu^{-2\epsilon} S_\epsilon \right)^2 \left(\frac{1}{v^B} \mathfrak{M}^B_{\text{SM}} + C_1^B \left(v^B\right)^2 \mathfrak{M}_1^B + C_2^B v^B \mathfrak{M}_2^B + C_3^B \mathfrak{M}_3^B \right), \end{aligned} \tag{5.10}$$

where α_s^B is the bare strong coupling parameter, m_t^B is the bare top mass, v^B is the d-dimensional vacuum expectation value of the Higgs field, and C_i^B are the bare Wilson coefficients. We also define $S_\epsilon = \exp(-\epsilon\gamma)(4\pi)^\epsilon$, the usual $\overline{\text{MS}}$ factor with γ the Euler-Mascheroni constant and μ is the renormal-

Operator	Vertex	Feynman Rule
O_1	$\bar{t}tH$	$iC_1\frac{v^2}{\sqrt{2}}$
O_2	g_1g_2H	$-8ig_s^2vC_2H^{\mu_1\mu_2}\delta^{c_1,c_2}$
O_3	$\bar{t}_1t_2g_3$	$ig_sC_3v\frac{\sqrt{2}}{2}T^{a_3}_{i_1i_2}\,[\not{p}_3,\gamma^{\mu_3}]$
O_3	$\bar{t}_1t_2g_3H$	$ig_sC_3\frac{\sqrt{2}}{2}T^{a_3}_{i_1i_2}\,[\not{p}_3,\gamma^{\mu_3}]$
O_3	$\bar{t}_1t_2g_3g_4$	$ig_s^2C_3\frac{\sqrt{2}}{2}vf_{a_3a_4a_5}T^{a_5}_{i_1i_2}[\gamma^{\mu_4},\gamma^{\mu_3}]$
O_3	$\bar{t}_1t_2g_3g_4H$	$ig_s^2C_3\frac{\sqrt{2}}{2}f_{a_3a_4a_5}T^{a_5}_{i_1i_2}[\gamma^{\mu_4},\gamma^{\mu_3}]$

Table 5.1: The new Feynman rules introduced in `qgraf-xml-parser`. The index of each particle is defined to correspond to the indices and momenta in the Feynman rules. The gluon fusion operator contains the transverse tensor $H^{\mu_1\mu_2} = (p_1^{\mu_2}p_2^{\mu_1} - p_1p_2g^{\mu_1\mu_2})$

ization scale. In the subsequent sections and until Section 5.3, we will only work with bare parameters and omit the superscript.

5.2.2 LO amplitude

The evaluation of the known result for the LO amplitude for gluon fusion in the SMEFT is a necessary step for the calculation of the NLO corrections. Not only does this amplitude provide a check of our implementation, but it is necessary for the renormalization and the subtraction of infrared divergences. Similarly, our choice of subtracting Higgs vacuum expectation values from the operators has a consequence on the SM renormalization in that the bare strong coupling constant and the bare top mass will include counterterms proportional to C_3. As a result, we need to compute the complete LO amplitude, including the SM contribution.

We generated diagrams using `QGRAF` to generate diagrams in `XML` format. We defined the new interactions generated by the dimension-six operators in `qgraf-xml-parser` to generate `FORM` expressions for $\mathcal{M}$ which could be processed to evaluate fermion traces and index contractions. The new Feynman rules generated by operators O_1, O_2 and O_3 are listed in Table 5.1.

One-loop diagrams

The integrals in the one-loop diagrams generated in the SM and by O_1 and O_3 all fall in the same triangle topology. We defined the following set of denominators:

$$D_1 = k_1^2 - m_t^2, \quad D_2 = (k_1 - p_1)^2 - m_t^2, \quad D_3 = (k_1 + p_2)^2 - m_t^2, \tag{5.11}$$

where k_1 is the loop momentum. Numerators are handled with the following rules:

$$k_1^2 = D_1 + m_t^2, \quad k_1 \cdot p_1 = \frac{1}{2}\left(D_1 - D_2\right), \quad k_1 \cdot p_2 = \frac{1}{2}\left(D_3 - D_1\right). \tag{5.12}$$

Integrals in the family are defined as

$$\mathcal{I}(n_1, n_2, n_3) = \mu^{2\epsilon} S_\epsilon^{-1} \int \frac{d^d k_1}{(2\pi)^d} \frac{1}{D_1^{n_1} D_2^{n_2} D_3^{n_3}}. \tag{5.13}$$

We used `SAIF` to reduce all the scalar integrals appearing in the one-loop diagrams into linear combinations of $\mathcal{I}$ integrals and found the following expressions:

- SM contribution:

$$\begin{aligned} M_{\text{SM}} =& 4\pi^2 \frac{m_t^2}{(d-2)s} \left(4\,\mathcal{I}(1,1,1)\left((d-2)s - 8m_t^2\right) - \frac{16\,\mathcal{I}(-1,1,1)}{s} \right. \\ &\left. + \frac{16\,\mathcal{I}(0,0,1)}{s} + \frac{16\,\mathcal{I}(0,1,0)}{s} - \frac{16\,\mathcal{I}(1,0,0)}{s} + 8(d-5)\,\mathcal{I}(0,1,1) \right) \end{aligned} \tag{5.14}$$

- Contribution from O_1:

$$\begin{aligned} M_1 =& 4\pi^2 \frac{m_t}{\sqrt{2}(d-2)s} \left(4\,\mathcal{I}(1,1,1)\left((d-2)s - 8m_t^2\right) - \frac{16\,\mathcal{I}(-1,1,1)}{s} \right. \\ &\left. + \frac{16\,\mathcal{I}(0,0,1)}{s} + \frac{16\,\mathcal{I}(0,1,0)}{s} - \frac{16\,\mathcal{I}(1,0,0)}{s} + 8(d-5)\,\mathcal{I}(0,1,1) \right) \end{aligned} \tag{5.15}$$

- Contribution from O_3

$$\begin{aligned} M_3 =& 4\pi^2 \frac{\sqrt{2} m_t}{(d-2)s} \left(16(d-3)sJ(1,1,1)m_t^2 + 8(d-4)sJ(0,1,1) \right. \\ &+ 2(d-2)sJ(1,0,1) + 2(d-2)sJ(1,1,0) - 4(d-6)J(0,0,1) \\ &\left. -4(d-6)J(0,1,0) + 8(d-4)J(1,0,0) - 16J(-1,1,1) \right) \end{aligned} \tag{5.16}$$

We reduced the scalar integrals appearing in these amplitudes to master integrals using `LiteRed` and obtained the expected tadpole ($\mathcal{I}(0,0,1)$), bubble ($\mathcal{I}(0,1,1)$) and triangle ($\mathcal{I}(1,1,1)$).

These integrals are of course well known and we used their expressions provided in [51]. These integrals are expressed in terms of multiple polylogarithm of the variable x defined by $s = -m_t^2(1-x)^2/x$. In the Euclidean region in which $s < 0$, the expression for x is

$$x = \frac{\sqrt{4m_t^2 - s} - \sqrt{-s}}{\sqrt{4m_t^2 - s} + \sqrt{-s}}. \tag{5.17}$$

We will discuss the analytic continuation of x and of the MPLs to the physical region in Section 5.5.

We can now express the one-loop diagrams appearing at LO in our calculation. We obtained each piece of the amplitude up to $\mathcal{O}(\epsilon^2)$, as needed to calculate the IR divergences of the NLO amplitude in

the Catani-Seymour formalism [188]. Here we present the result up to $\mathcal{O}(\epsilon^0)$:

$$M_{\text{SM}}^B = -i\frac{x\left((x+1)^2 G(0,0,x) - 2(x-1)^2\right)}{(x-1)^4} \tag{5.18}$$

$$M_1^B = -\frac{i}{\sqrt{2}m_t}\frac{x\left((x+1)^2 G(0,0,x) - 2(x-1)^2\right)}{(x-1)^4} \tag{5.19}$$

$$\begin{aligned} M_3^B = {} & \frac{i2\sqrt{2}m_t}{\epsilon} \\ & - \frac{i\sqrt{2}m_t\left(\left(x^2-1\right)G(0,x) + 2xG(0,0,x) - (x-1)^2\left(\left(2\log\left(\frac{\mu^2}{m_t^2}\right)+1\right)\right)\right)}{(x-1)^2} \end{aligned} \tag{5.20}$$

As we discussed in Section 5.1, the diagrams generated by O_3 have a UV divergence. The calculation of the tree-level contribution from O_2 will show that it can be absorbed by the renormalization of C_2.

Tree-level diagram

The trivial contribution generated by O_2 is very simple to calculate. Applying the relevant Feynman rule in Table 5.1, we find

$$M_2^B = -32i\pi^2. \tag{5.21}$$

As we discussed earlier, C_2^B contains a counterterm $C_3^R \delta Z_{23}/\epsilon$, which cancels the pole in Equation 5.20 by setting

$$\delta Z_{23} = \frac{m_t}{8\sqrt{2}\pi^2 v}, \tag{5.22}$$

in agreement with the literature [112, 114, 115]. We will discuss in detail in Section 5.3 the full one-loop renormalization matrix for our set of three operators.

5.2.3 NLO amplitude

We will now calculate the $\mathcal{O}(\alpha_s^2)$ virtual corrections to gluon fusion generated by the SM and our set of three operators. The SM diagrams and those generated by O_1 and O_3 are two-loop diagrams, and the diagrams generated by O_2 are one-loop diagrams which we will handle separately.

Using the now standard chain of tools, we generated independently the two-loop and the one-loop diagrams using `QGRAF`, which we then transformed to `FORM` expressions for $\mathcal{M}_{\text{SM}}$, $\mathcal{M}_i$, using `qgraf-xml-parser` whose indices were contracted and gamma matrix traces evaluated. The resulting expressions were then imported into `Mathematica` for further processing.

One-loop diagrams

There is only one one-loop diagrams generated by O_2 at one loop, as shown in Figure 5.2. It is therefore straightforward to define a topology to evaluate the loop integral in this diagram. Let us define the

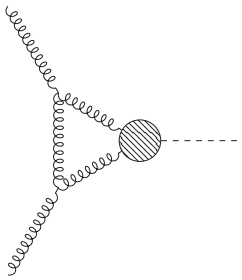


Figure 5.2: The one-loop diagram contributing to $\mathcal{M}_2^B$

family

$$j(n_1, n_2, n_3) = S_\epsilon^{-1} \mu^{2\epsilon} \int \frac{d^d k_1}{(2\pi)^d} \frac{1}{(k_1^2)^{n_1} ((k_1 - p_1)^2)^{n_2} ((k_1 + p_2)^2)^{n_3}} \tag{5.23}$$

We used `SAIF` to reduce all the scalar integrals appearing in $\mathcal{M}_2^B$ to linear combinations of j integrals. We then used `LiteRed` to reduce these integrals to a basis of only one master integral, $j(0, 1, 1)$, a massless bubble. This integral is trivial to evaluate using the techniques described in Section 4, but in practice we simply used the expression provided in [51], as we had already imported their definitions for the two-loop part of our calculation. Inserting the expression for $j(0, 1, 1)$ into $\mathcal{M}_2^B$, we get a somewhat complicated expression which hides the actual simplicity of the amplitude:

$$\begin{aligned} \mathcal{M}_2^B = 16i\pi^2 \Bigg[& \frac{3}{\epsilon^2} + \frac{1}{\epsilon} \left(3G(0, x) - 6G(1, x) + \frac{\beta_0}{2} + 3 \log \left(\frac{\mu^2}{m_t^2} \right) \right) \\ & + 3G(0, 0, x) - 6G(0, 1, x) - 6G(1, 0, x) + 12G(1, 1, x) - \frac{\pi^2}{4} \\ & + 3G(0, x) \log \left(\frac{\mu^2}{m_t^2} \right) - 6G(1, x) \log \left(\frac{\mu^2}{m_t^2} \right) + \frac{3}{2} \log^2 \left(\frac{\mu^2}{m_t^2} \right) \Bigg] \end{aligned} \tag{5.24}$$

This is due to the fact that our choice of variables (x, m_t) is not well-suited to this result, which is entirely independent of m_t. We can however simplify this expression greatly by reverting to s. To this end, we can use the fact that harmonic polylogarithms have known values at $x = 1/2$ which can be computed by `PolyLogTools` using the technique proposed in [160]. We can evaluate each polylogarithm in our expression in terms of ζ-values and $\log 2$ and use the fact that, at this point, $m_t^2 = -2s$ to express $\mathcal{M}_2^B$ in terms of s. We find the much simpler result

$$\mathcal{M}_2^B = 16i\pi^2 \left(\frac{3}{\epsilon^2} + \frac{1}{\epsilon} \left(\frac{\beta_0}{2} + 3 \log \left(-\frac{\mu^2}{s} \right) \right) + \frac{3}{2} \log^2 \left(-\frac{\mu^2}{s} \right) - \frac{\pi^2}{4} \right), \tag{5.25}$$

which we checked numerically against Equation 5.24.

Two-loop diagrams

The calculation of the two-loop diagrams generated by the SM and the operators O_1 and O_3 is significantly more complex than the evaluation of the one-loop diagrams generated by O_2, in part because there are many more diagrams, which require several different topologies.

It is easy to see that the set of topologies defined in the SM calculation of Higgs production at

NLO [51, 52] is sufficient for all the integrals we need to evaluate. Indeed, in our calculation, O_1 only brings a correction to the value of the $Ht\bar{t}$ vertex, meaning that there is a one-to-one matching of the SM diagrams to those generated by O_1. Furthermore, of the four vertices generated by O_3, one is already present in the SM, and three can be viewed as pinchings of SM diagrams, as shown in Figure 5.3. Therefore, we use the same three families as [51], which we call T_1, T_2, T_3, and whose integrals are

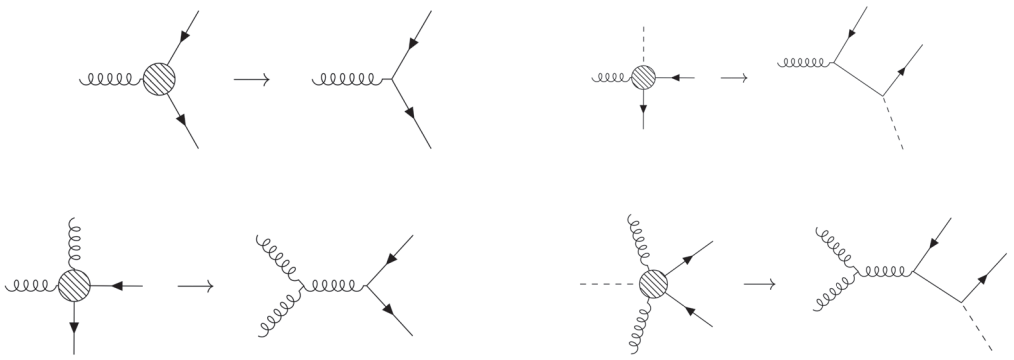


Figure 5.3: The vertices introduced by O_3 can be seen as (pinched) SM interactions.

$$\mathcal{J}_{T_i}(n_1,\ldots,n_7) = S_\epsilon^{-2}\mu^{4\epsilon} m_t^{4\epsilon+2\sum n_k} \int \frac{d^dk_1}{(2\pi)^d}\frac{d^dk_2}{(2\pi)^d}\frac{1}{\left(D_1^{(i)}\right)^{n_1}\cdots\left(D_7^{(i)}\right)^{n_1}}, \tag{5.26}$$

where the factors of m_t ensure that the integrals don't have a mass dimension, which will make the IBP relations simpler, and the denominators $D_j^{(i)}$ are:

T_i	T_1	T_2	T_3
$D_1^{(i)}$	k_1^2	$k_1^2 - m_t^2$	$k_1^2 - m_t^2$
$D_2^{(i)}$	$(k_1+p_1)^2$	$(k_1+p_2)^2 - m_t^2$	$(k_1-k_2-p_1)^2$
$D_3^{(i)}$	$(k_1+p_1+p_2)^2$	$(k_1+p_1+p_2)^2 - m_t^2$	$(k_1+p_1+p_2)^2 - m_t^2$
$D_4^{(i)}$	$(k_2+p_1+p_2)^2 - m_t^2$	$(k_2+p_1+p_2)^2 - m_t^2$	$(k_2+p_1+p_2)^2 - m_t^2$
$D_5^{(i)}$	$(k_2+p_1)^2 - m_t^2$	$(k_2+p_2)^2 - m_t^2$	$(k_2+p_1)^2 - m_t^2$
$D_6^{(i)}$	$k_2^2 - m_t^2$	$k_2^2 - m_t^2$	$(k_1+p_1)^2 - m_t^2$
$D_7^{(i)}$	$(k_1-k_2)^2 - m_t^2$	$(k_1-k_2)^2$	$(k_1-k_2)^2$.

(5.27)

The fact that we can use the same topologies as the existing SM calculations provides a tremendous simplification of the calculation of our two-loop diagrams. Indeed, as a result, all the master integrals in each topology are known analytically and were computed in [51], and in [52]. We decided to use the results from [52] for all integrals that do not factorize into products of one-loop integrals for very practical purposes: they provided their results in the form of electronic files which were straightforward to convert to `Mathematica`. Since we had already implemented the one-loop integrals provided by [51], we used these results for the factorized master integrals. The set of non-factorized integrals whose explicit values we use is denoted Int_a...Int_k and shown in Figure 5.4. We provide the explicit definition of the integrals with numerators as momentum-space integrals in Appendix D.1. The factorized master integrals are denoted $\text{Int}_{xy} = \text{Int}_x\text{Int}_y$, where x, y are picked from {b, mb, t, tri}, which correspond

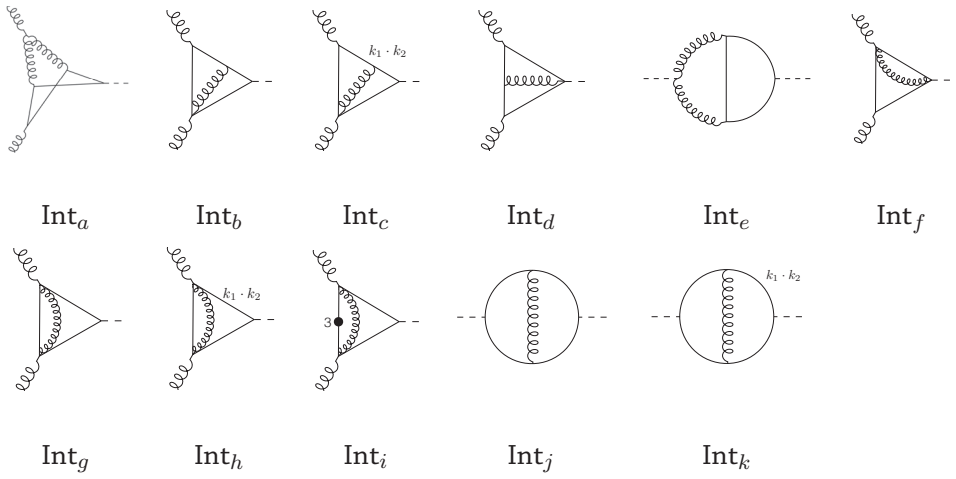


Figure 5.4: The set of master integrals from [52].

respectively to a massless bubble integral, a massive bubble, a tadpole integral, and a massive triangle integral. Each of these integrals has been obtained as a series in ϵ and each term in the series is a linear combination of harmonic polylogarithms of x.

We used `SAIF` to sort all the scalar integrals in the 21 SM diagrams, the 21 O_1 diagrams and the 75 O_3 diagrams. In total, we found that $\mathcal{M}^B_{\text{SM}}$, $\mathcal{M}^B_1$ and $\mathcal{M}^B_3$ can be expressed in terms of 821 T_1 integrals, 635 T_2 integrals and 513 T_3 integrals. We then used `LiteRed` to express these integrals in terms of master integrals for each topology and insert these expressions into $\mathcal{M}^B_{\text{SM}}$, $\mathcal{M}^B_1$ and $\mathcal{M}^B_3$. As could be expected, the reduction in `LiteRed` resulted in a different set of master integrals than that of [52]. We therefore needed to express each master integral found by `LiteRed` as a linear combination of known integrals.

To this end, we used `SAIF` to sort the 18 masters from the literature into our topologies. We expressed each known integral as an explicit momentum space integral and tried to sort it into T_1, T_2 and T_3. Because there is some overlap between the bases of each family, some integrals were matched to more than one family, but for each family the number of matching integrals was exactly that of master integrals in the basis found by `LiteRed`. For each family, we could then reduce the integrals from the literature as linear combinations of the `LiteRed` bases to generate a linear system equations which we could then solve to find the expression of the three `LiteRed` bases in terms of the known integrals.

Using the solution of this system, we could express our scalar amplitudes $\mathcal{M}_{\text{SM}}$, $\mathcal{M}_1$ and $\mathcal{M}_3$ as a series in ϵ up to order ϵ^0 where each term is a linear combinations of harmonic polylogarithms of x. We checked that our implementation of the analytic formulae found in the literature and our change of basis to the master integrals defined by `LiteRed` were correct numerically. We evaluated directly each of the integrals in the bases defined by `LiteRed` (see Appendix D.2 for the list of integrals) using `FIESTA` and `SecDec` and compared this result with the evaluation of these integrals expressed as linear combinations of harmonic polylogarithms using `Ginac` and found good agreement for several choices of x and m_t both in the Euclidean ($s < 0$) and in the physical ($4m_t^2 > s > 0$) regions. The expressions for $\mathcal{M}_1$ and $\mathcal{M}_3$ are provided in Appendix D.4.

5.3 Structure of UV and IR divergences

In this section, we discuss the divergence structure of our NLO amplitude. We will first describe the 1-loop and 2-loop UV divergences that appear in our calculation and then discuss the IR structure of the NLO amplitude.

For the renormalization of UV divergences, we use a modified $\overline{\text{MS}}$ scheme for the parameters, setup so that top quark and dimension-six operator effects decouple from SM parameters [189]. The precise set of rules governing this scheme is layed out in Appendix D.3.

The relation between the bare amplitude $\mathcal{M}$ and the renormalized amplitude $\mathcal{M}^{\mathcal{R}}$ is:

$$Z_g \mathcal{M}(\alpha_s^B, m_t^B, C_i^B, v^B) = \mathcal{M}^R(\alpha_s^R, m_t^R, C_i^R, v^R), \tag{5.28}$$

where Z_g is the gluon field renormalization constant, α_s^R, m_t^R and C_i^R are the renormalized strong coupling constant, the renormalized top mass and the renormalized Wilson coefficients, which we will assume to be evaluated at the renormalization scale μ. As in the previous chapter, we must separate the d-dimensional Higgs vacuum expectation value v^B from the 4-dimensional parameter $v^R = \mu^\epsilon v^B$ to preserve the homogeneity of the amplitude.

The renormalized parameters are related to the bare parameters by

$$\alpha_s^B = S_\epsilon^{-1} \mu^{2\epsilon} Z_\alpha \alpha_s^R, \tag{5.29}$$

$$m_t^B = Z_m m_t^R, \tag{5.30}$$

$$C_i^B = \mu^{\kappa_i \epsilon} \left(Z_C\right)_{ij} C_j^R, \tag{5.31}$$

where $\vec{\kappa} = (3, 0, 1)$.

The one-loop UV counterterms for the strong coupling constant and the gluon field are given by [118]

$$\begin{aligned} Z_g &= 1 + \delta Z_{g,SM} + \frac{\alpha_s}{\pi} C_3 \frac{1}{\epsilon} \left(\frac{\mu^2}{m_t^2}\right)^\epsilon \sqrt{2}\, v\, m_t + \mathcal{O}(\alpha_s^2) \ , \\ Z_{\alpha_s} &= 1 + \delta Z_{\alpha_s,SM} - \frac{\alpha_s}{\pi} C_3 \frac{1}{\epsilon} \left(\frac{\mu^2}{m_t^2}\right)^\epsilon \sqrt{2}\, v\, m_t + \mathcal{O}(\alpha_s^2) \ , \end{aligned} \tag{5.32}$$

where $\delta Z_{g,SM}$ and $\delta Z_{\alpha_s,SM}$ denote the one-loop UV counterterms in the SM,

$$\begin{aligned} \delta Z_{g,SM} &= \frac{\alpha_s}{\pi} \frac{1}{6\epsilon} \left(\frac{\mu^2}{m_t^2}\right)^\epsilon + \mathcal{O}(\alpha_s^2) \ , \\ \delta Z_{\alpha_s,SM} &= -\frac{\alpha_s}{4\pi} \frac{\beta_0}{\epsilon} - \frac{\alpha_s}{\pi} \frac{1}{6\epsilon} \left(\frac{\mu^2}{m_t^2}\right)^\epsilon + \mathcal{O}(\alpha_s^2) \ , \end{aligned} \tag{5.33}$$

where β_0 is the one-loop QCD β function,

$$\beta_0 = \frac{11 N_c}{3} - \frac{2}{3} N_f \, , \tag{5.34}$$

where $N_c = 3$ is the number of colours and $N_f = 5$ is the number of massless flavours. Our choice to decouple the top quark from the running of the strong coupling constant shows in the factors of

$(\mu^2/m_t^2)^\epsilon$ that appears in the terms coming from top loops, whose effect is to cancel top quark effects in the running of α_s.

Similarly, the renormalisation of the top mass is modified by the presence of the effective operators [118],

$$Z_m = 1 + \delta Z_m^{SM} - \frac{\alpha_s}{\pi} C_3 \frac{1}{\epsilon} \left(\frac{\mu^2}{m_t^2}\right)^\epsilon 2\sqrt{2}\, v\, m_t + \mathcal{O}\big(\alpha_s^2\big) \ , \tag{5.35}$$

where the SM constribution is

$$\delta Z_m^{SM} = -\frac{\alpha_s}{\pi} \frac{1}{\epsilon} + \mathcal{O}\big(\alpha_s^2\big) \ . \tag{5.36}$$

Note that in Equation 5.35 we also include the $\left(\mu^2/m_t^2\right)^\epsilon$ into the counterterm in order to decouple the effects from operators of dimension six from the running of the top mass.

The renormalisation of the effective couplings C_i^B is more involved, because the Wilson coefficients mix under renormalisation. The matrix Z_C of counterterms can be written in the form

$$Z_C = \mathbb{1} + \delta Z_C^{(0)} + \frac{\alpha_s}{\pi} \delta Z_C^{(1)} + \mathcal{O}\big(\alpha_s^2\big) \ . \tag{5.37}$$

We have already mentioned that the amplitude M_3 requires renormalisation despite contributing at leading-order in the strong coupling, and the UV divergence is proportional to the leading order amplitude M_2, as was illustrated in Equation 5.7. As a consequence, $\delta Z_C^{(0)}$ is non-trivial:

$$\delta Z_C^{(0)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \dfrac{\sqrt{2}\, m_t}{16\pi^2\, \epsilon\, v} \\ 0 & 0 & 0 \end{pmatrix} \ . \tag{5.38}$$

At NLO, we also need the contribution $\delta Z_C^{(1)}$ to eq. (5.37). We have [112, 114, 115]

$$\delta Z_C^{(1)} = \begin{pmatrix} -\dfrac{1}{\epsilon} & 0 & -\dfrac{8{m_t}^2}{\epsilon\, v^2} \\ 0 & 0 & z_{23} \\ 0 & 0 & \dfrac{1}{6\,\epsilon} \end{pmatrix} \ . \tag{5.39}$$

The quantity z_{23} denotes the counterterm that absorbs the two-loop UV divergence of the operator O_3, which is proportional to the tree-level amplitude M_2 in our case. This counterterm is not available in the literature, but we can extract it from our computation. In order to do so, we need to construct a quantity that is finite in the limit $\epsilon \to 0$, which our renormalized amplitude at NLO is not because it contains IR divergences. By subtracting these infrared divergences from the amplitude, we construct a finite remainder that still has a dependence on z_{23}, whose value we can fix by imposing that all divergences cancel.

If we consider a renormalized amplitude $\mathcal{A}^{\mathcal{R}} = (\alpha_s/\pi)^k \left(M^R + (\alpha_s/\pi)\, \mathcal{M}^R\right)$, the NLO correction to the amplitude has IR divergences which take the form [188]

$$\mathcal{M}^R = \mathcal{M}_F^R + I_1 M^R, \tag{5.40}$$

where $\mathcal{M}_F^R$ is finite and I_1 is the universal dipole operator that depends only on the external kinematics and color (and not on the details of the hard scattering). In general,

$$I_1 = \frac{1}{2}\frac{e^{\epsilon\gamma}}{\Gamma(1-\epsilon)}\sum_i \frac{1}{\mathbf{T}_i^2}\left(\mathbf{T}_i^2\frac{1}{\epsilon^2}+\gamma_i\,\frac{1}{\epsilon}\right)\sum_{j\neq i}\mathbf{T}_i\cdot\mathbf{T}_j\,\left(\frac{\mu^2}{-2\,p_i\cdot p_j}\right)^{\epsilon}, \tag{5.41}$$

where the sums run over the colored external states of the amplitude, the T_i are the generator of the representation of the corresponding parton and

$$\gamma_q = \gamma_{\bar{q}} = \tfrac{3}{2}\,C_F, \quad \gamma_g = \tfrac{11}{6}\,C_A - \tfrac{1}{3}\,N_f, \tag{5.42}$$

respectively for quarks, anti-quarks and gluons. In the case of the production of a colourless state from the scattering of two gluons, I_1 simplifies to

$$I_1 = -\frac{1}{2}\frac{e^{-\gamma\epsilon}}{\Gamma(1-\epsilon)}\left(\frac{-s_{12}}{\mu^2}\right)^{-\epsilon}\left(\frac{3}{\epsilon^2}+\frac{\beta_0}{2\epsilon}\right). \tag{5.43}$$

We subtracted the IR divergences from the NLO amplitudes in our calculations to build

$$\mathcal{M}_{F,\text{SM}}^R = \mathcal{M}_{\text{SM}}^R - I_1 M_{\text{SM}}^R \tag{5.44}$$

$$\mathcal{M}_{F,i}^R = \mathcal{M}_i^R - I_1 M_i^R \;\; i = 1,\ldots,3, \tag{5.45}$$

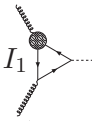
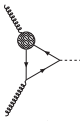
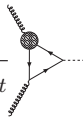
out of the amplitudes we have calculated in the previous section. We checked for all the amplitudes but $\mathcal{M}_3^R$ that they had the correct IR divergence structure and that, indeed, the subtracted remainders $\mathcal{M}_{F,\text{SM}}^R$, $\mathcal{M}_{F,1}^R$ and $\mathcal{M}_{F,2}^R$ are finite in the $\epsilon \to 0$ limit. On $\mathcal{M}_{F,3}^R$, we used Equation 5.45 as a constraint to fix the value of z_{23} by imposing that all poles cancel. We find

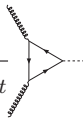
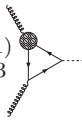
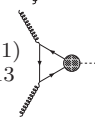
$$z_{23} = \frac{m_t}{16\,\pi^2\,v\,\sqrt{2}}\left(-\frac{5}{6\,\epsilon^2}+\frac{23}{4\,\epsilon}\right). \tag{5.46}$$

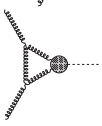
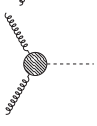
The value of the double pole is in principle fixed by imposing that the $\mathcal{O}(\alpha_s)$ anomalous dimension matrix for our Wilson coefficients is finite. We have checked that this criterion is satisfied by our result, which provides a strong check of our calculation.

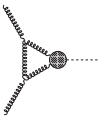
A remark about the divergence structure of $\mathcal{M}_3$ The folk wisdom about QCD radiative corrections tells us what to expect of the degree of divergence of NLO virtual corrections to amplitudes. Indeed, because in the Standard Model, LO calculations are always UV finite, we know from the renormalization structure of the amplitude that the NLO corrections will contain $1/\epsilon$ poles of UV and IR origin and $1/\epsilon^2$ poles of IR origin. However, in our SMEFT calculation, this picture is perturbed by the fact that we can freely assign the scaling in α_s of the Wilson coefficients, which in our case leads to UV divergences at LO. This is not an issue because our choice also ensured that the divergences in the LO loop-induced diagrams can be absorbed by the tree-level diagram. Once we work with renormalized amplitudes, the structure of IR divergences described in [188] is respected. This is easily understandable from the fact that the remaining IR divergences in the NLO correction to the amplitude after renormalization have to be cancelled by the real emissions, which entirely factorize from the hard scattering and

therefore are agnostic of whether the LO diagrams are tree-level or one-loop, or if divergences appear in the UV. We found indeed that our renormalized amplitudes have at most $1/\epsilon^2$ poles and that they have exactly the structure predicted by Equation 5.40. It is however interesting to look at the divergences of the bare amplitude. The bare amplitude generated by O_3, $\mathcal{M}_3$ contains poles up to $1/\epsilon^3$. Let us schematically decompose $\mathcal{M}_3$ in terms of its UV and IR divergences to understand their origin:

$\mathcal{M}_3 = \mathcal{M}^R_{3,F} + I_1$  $- \delta Z^{\text{SM}}_{\alpha_s}$  $- m_t \delta Z^{\text{SM}}_m \frac{\partial}{\partial m_t}$ 

$-m_t \delta Z^C_m \frac{\partial}{\partial m_t}$  $- (\delta Z_C)^{(1)}_{33}$  $- (\delta Z_C)^{(1)}_{13}$  (5.47)

$- (\delta Z_C)^{(0)}_{23}$  $- (\delta Z_C)^{(1)}_{23}$ 

The source of the $1/\epsilon^3$ divergence is the term $(\delta Z_C)^{(0)}_{23}$  $= (\delta Z_C)^{(0)}_{23}\, \mathcal{M}_2$. Indeed, the counterterm $(\delta Z_C)^{(0)}_{23}$ contains a $1/\epsilon$ pole, while the amplitude $\mathcal{M}_2$ verifies

$$\mathcal{M}_2 = \mathcal{M}^R_2 + (1/\epsilon \text{ UV poles}) = \mathcal{M}^R_{2,F} + I_1 M^R_2 + (1/\epsilon \text{ UV poles}). \tag{5.48}$$

Because I_1 contains $1/\epsilon^2$ poles, the overal divergence of $(\delta Z_C)^{(0)}_{23}\, \mathcal{M}_2$ is $1/\epsilon^3$. This is a result of the mixing of loop orders and of the LO renormalization of the amplitude. It still fits the general picture of UV and IR divergences of NLO amplitudes perfectly, but contradicts the common wisdom that they should at most contain $1/\epsilon^2$ poles because this expectation is based on the fact that LO amplitudes are usually UV finite.

5.4 Renormalization group evolution of the parameters

Our renormalized amplitude depends explicitly on the renormalization scale μ, which we identify with the factorization scale μ_F which will be used in the PDF for the evaluation of the hadronic cross-section. It will however be useful to understand the evolution of our parameters when we change the scale at which they are evaluated, first because it might be desirable to evaluate the cross section with different choices of scales for the strong parameter, the top mass and the Wilson coefficients, but also because it will allow us to evaluate the scale dependence of our result to give an estimate of the importance of missing logarithms at higher orders - which is our best guess for the importance of the next QCD corrections.

Because we work in a decoupling scheme, neither the running of the strong coupling constant nor that of the top mass differ from the SM case with 5 flavors in $\overline{\text{MS}}$. We have checked explicitly that this is indeed the case. Note that in our case we need to take into account the running of the Higgs vacuum expectation value v, which vanishes in the limit $\epsilon \to 0$. We find the renormalization group equation

(RGE) for v using the fact that the bare parameter does not depend on μ

$$0 = \frac{\partial v^B}{\partial \log \mu^2} = \frac{\partial v^R \mu^{-\epsilon}}{\partial \log \mu^2} \tag{5.49}$$

$$\implies \frac{\partial v^R}{\partial \log \mu^2} = \frac{1}{2}\epsilon v^R. \tag{5.50}$$

The extraction of the RGE for the other parameters is obtained in the same manner. We exploited the fact that none of the bare parameters depend on μ to generate a coupled system of differential equations that involve α_s^R, m_t^R, and the Wilson coefficients C_i. Our scheme choice ensures however that up to order $\mathcal{O}(\alpha_s)$, neither the derivative of α_s or that of m_t depend on the derivatives of the Wilson coefficients – while the reverse is of course true. Solving these equations in an expansion in ϵ and α_s, find that in the limit $\epsilon \to 0$

$$C_1(\mu^2) = C_1(Q^2) - \frac{\alpha_s(Q^2)}{\pi} \log \frac{\mu^2}{Q^2} \left(C_1(Q^2) + 8\, C_3(Q^2)\, \frac{m_t^2(Q^2)}{v^2} \right) + \mathcal{O}\big(\alpha_s(Q^2)^2\big) \ , \tag{5.51}$$

$$C_2(\mu^2) = C_2(Q^2) + \sqrt{2}\, \frac{C_3(Q^2)}{16\,\pi^2} \log \frac{\mu^2}{Q^2} \frac{m_t(Q^2)}{v} \tag{5.52}$$

$$- \sqrt{2}\, \frac{\alpha_s(Q^2)}{192\,\pi^3}\, C_3(Q^2) \log \frac{\mu^2}{Q^2} \frac{m_t(Q^2)}{v} \left(5 \log \frac{\mu^2}{Q^2} - 69 \right) + \mathcal{O}\big(\alpha_s(Q^2)^2\big) \ , \tag{5.53}$$

$$C_3(\mu^2) = C_3(Q^2) + C_3(Q^2)\, \frac{\alpha_s(Q^2)}{6\pi} \log \frac{\mu^2}{Q^2} + \mathcal{O}\big(\alpha_s(Q^2)^2\big) \ . \tag{5.54}$$

Note that C_1 and C_3 are renormalized at one loop and therefore only contain a simple logarithm, while C_2 is renormalized at two loops, meaning that it contains a simple logarithm in the first order term in its evolution, and a squared logarithm as well as a simple logarithm in the second order correction.

5.5 Analytic continuation of the amplitude to the physical regions

Up to this point, the entirety of our calculation has been performed in the Euclidean region in which $s < 0$, which we can translate into $0 < x < 1$. In this section, we will discuss the analytic continuation of the amplitude to the physical regions. We have used the top quark by default throughout this chapter, but in this section we will discuss the general case of the analytic continuation, which depends on whether the Higgs mass is above or below the quark pair production threshold. As a result, we will revert to discussing a generic quark of mass m running in the loop. The relation between x and s in this region is

$$x = \frac{\sqrt{4m^2 - s} - \sqrt{-s}}{\sqrt{4m^2 - s} + \sqrt{-s}}, \tag{5.55}$$

and our amplitude is expressed in terms of functions of x that can have branch cuts. As a result, we first need to understand the analytic continuation of $x(s)$ and of functions of x_s to the regions

$$\begin{aligned} R_1 &: 4m^2 > s > 0 \\ R_2 &: s > 4m^2. \end{aligned} \tag{5.56}$$

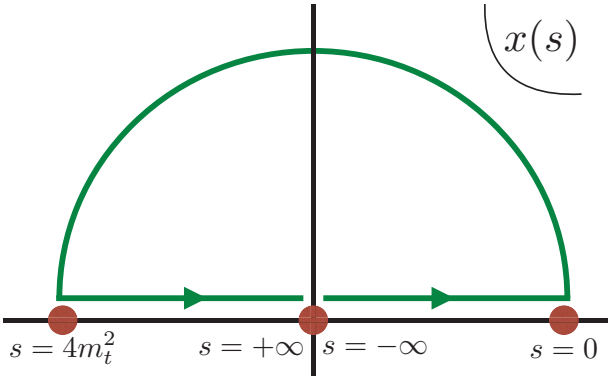


Figure 5.5: The path of $x(s)$ in $\mathbb{C}$ for s going from $-\infty$ to $+\infty$.

The region R_1 is relevant for calculations of the amplitude in which the Higgs mass is below twice the quark mass, *i.e.* for top loops, and R_2 is relevant for the other quark flavors.

For any multivalued function $F(s)$ which has a branch cut starting at some $s_0 \in \mathbb{R}^+$, the physical analytic continuation along the branch is obtained by approaching the limit by the upper half-plane [190]

$$F_{\text{physical}}(s)\big|_{s\in\mathbb{R}^+} = \lim_{\varepsilon\to 0} F(s + i\epsilon). \tag{5.57}$$

In particular, $\log(-s)$ is analytically continued to $\log s - i\pi$ for $s > 0$ and $\sqrt{-s}$ is continued to $-i\sqrt{s}$. In our case, it is now easy to find the analytic continuation for x in R_1:

$$x(s)|_{4m^2>s>0} = \frac{\sqrt{4m^2 - s} + i\sqrt{s}}{\sqrt{4m^2 - s} - i\sqrt{s}}, \tag{5.58}$$

which means that in R_1, $x(s)$ is on the upper-half complex plane. In the region above threshold R_2,

$$x(s)|_{s>4m^2} = \frac{\sqrt{s - 4m^2} - \sqrt{s}}{\sqrt{s - 4m^2} + \sqrt{s}}, \tag{5.59}$$

and as a result $-1 < x < 0$. The path of x in the complex plane as s goes from $-\infty$ to $+\infty$ is described in Figure 5.5

Analytically continuing x is not the end of the story. Indeed, our amplitude is expressed in terms of harmonic polylogarithms of x, which can have branch cuts. The possible branch cuts of a polylogarithm $G(a_1, \ldots, a_n, x)$ start at the points $a_i \in \{-1, 0, 1\}$. Branch cuts starting at $x = 0$ and $x = -1$ extend to $x = -\infty$ and those starting at $x = 1$ extend to $x = +\infty$ [160]. As a result, two types of issues can arise:

- individual polylogarithms in the amplitude can have divergences at two physical points: $x = 1$ and $x = -1$. The amplitude is of course finite and all such divergences have to cancel against each other. If we wanted to evaluate our amplitude at either point, we would need to extract the relevant logarithmic divergences from each function and cancel them explicitly. As the two physical regions of interest are not located at these points, we do not need to worry about this issue.

- Evaluating the amplitude in R_2, however, means evaluating some polylogarithms along a branch cut since we have $x < 0$. Let us discuss how to solve this issue below. Note that because there is no branch cut on the upper unit circle, we can evaluate the amplitude in R_1 by using the value for x given by Equation 5.58.

We know that harmonic polylogarithms whose last entry is not a 0 do not have a branch cut along $\mathbb{R}^-$. In such a case, we have the relation

$$G(a_1, \ldots, a_n, -x) = (-1)^n G(-a_1, \ldots, -a_n, x) \qquad (\text{if } a_n \neq 0). \tag{5.60}$$

The condition that the amplitude must be analytic and the known path of x in the complex plane gives us a simple way to analytically continue simple logarithms. As shown in Figure 5.5, below threshold, x is a unit complex number in the upper half plane so $\log(x) = i\theta$ for some $\theta \in [0, \pi]$. In particular, $\lim_{s \to 4m^2} \log(x) = i\pi$ so the continuity of the amplitude imposes that the right branch to analytically continue $\log(x)$ to $-1 < x < 0$ is

$$\log(x)\big|_{-1<x<0} := \log(-x) + i\pi \tag{5.61}$$

Generalizing this to arbitrary functions $F(x)$ with a branch cut along $x < 0$, the physical branch of F above threshold is

$$F_{\text{phys}}(x)\big|_{-1<x<0} = \lim_{\varepsilon \to 0} F(x + i\epsilon). \tag{5.62}$$

While we could perform this procedure on each polylogarithm that has a 0 in its rightmost entry, there is a more algorithmic way of finding their analytic continuation above threshold. Indeed, we know how to factorize logarithmic divergences out of polylogarithms, *i.e.* how to recast any $G(\ldots, 0, x)$ into a linear combination of products of poylogarithms where the only functions with branch cuts along the line $x < 0$ are functions of the form $G(\underbrace{0, \ldots, 0}_{n}, x) = 1/n! \log^n x$, which we know how to analytically continue to R_2.

Algorithmically, this means that we transform our expression in three steps:

- Extract the $G(0, \ldots, 0, x)$ from the polylogarithms in the amplitude (this is done using the `PolyLogTools` function `ExtractZeroes`)
- Replace all $G(a_1, \ldots, \pm 1, x)$ by $G(-a_1, \ldots, \mp 1, -x)$
- Replace all $G(0, \ldots, 0, x)$ by

$$\frac{1}{n!}(\log(-x) + i\pi)^n = \sum_{k=0}^{n} \frac{(i\pi)^{n-k}}{(n-k)!} G(\underbrace{0, \ldots, 0}_{k}, -x). \tag{5.63}$$

After these operations, all the polylogarithms in the amplitude have $-x$ as argument and are therefore defined without ambiguities in the region R_2.

6

Gluon fusion in the SMEFT at NLO

6.1 Calculation setup

The results presented in the previous chapter can be used for a phenomenological study of Higgs production in the SMEFT. We combined the virtual corrections to gluon fusion obtained analytically with the real emission diagrams generated automatically with `MadGraph5_aMC@NLO`. We built on the implementation developped for the calculation of Higgs boson production in association with one jet at LO in the SMEFT [118], which is described by the same Feynman diagrams as the real emission corrections for inclusive gluon fusion at NLO. In `MadGraph5_aMC@NLO`, the virtual corrections and the real emissions are combined to evaluate the partonic cross section numerically using the Monte Carlo method, which is then integrated over the PDF of the proton. The calculation is matched to the parton shower generator `PYTHIA8` that deals with further parton emissions.

To match the scheme used in `MadGraph5_aMC@NLO`, we changed the mass scheme for the top quark to the on-shell scheme. Let us show how we can relate the mass we have used in our decoupling scheme $\overline{m_t}^{\text{dec}}(\mu)$ to the on-shell mass m_t^{os}. They are both related to the bare mass:

$$m_t^B = Z_m \overline{m_t}^{\text{dec}}(\mu) = \hat{Z}_m m_t^{\text{os}}, \tag{6.1}$$

where Z_m and $\hat{Z}_m$ are respectively the top mass renormalization constants in our scheme and in the on-shell scheme. Taking the value for $\hat{Z}_m$ from [118], we have

$$\overline{m_t}^{\text{dec}}(\mu) = m_t^{\text{os}} \frac{\hat{Z}_m}{Z_m} = m_t^{\text{os}} \left(1 - \frac{\alpha_s}{\pi}\left(\frac{4}{3} + \log\left(\frac{\mu^2}{m_t^2}\right) + C_3 \frac{2\sqrt{2}}{3} v m_t\right)\right). \tag{6.2}$$

The Standard Model results presented here have been cross-checked with the NLO+PS implementation of `aMCSusHi` [191]. Results are obtained with MMHT2014 LO/NLO PDFs [192], for LO and NLO

results respectively; input parameters are

$$m_t = 173\ \text{GeV}\,, \quad m_h = 125\ \text{GeV}\,, \quad m_Z = 91.1876\ \text{GeV}\,, \tag{6.3}$$

$$\alpha_{EW}^{-1} = 127.9\,, \quad G_F = 1.16637 \times 10^{-5}\ \text{GeV}^{-2}\,. \tag{6.4}$$

The factorization and renormalization scales are set at $m_h/2$. We introduce a new scale μ_{EFT} at which we evaluate the Wilson coefficients, also set at $\mu = m_h/2$, which we will use to assess uncertainties related to the running of these couplings, as discussed in the next section. The running of the Wilson coefficients is controlled by a running operator

$$C_i(\mu^2) = \Gamma_{ij}(\mu^2, \mu^2_{\text{EFT}}) C_j(\mu^2_{\text{EFT}}). \tag{6.5}$$

Cross sections from dimension-six operators can be parametrized as

$$\sigma(\mu) = \sigma_{SM}(\mu^2) + \sum_i \frac{1\text{TeV}^2}{\Lambda^2} C_i(\mu^2)\sigma_i(\mu^2) \tag{6.6}$$

$$= \sigma_{SM}(\mu^2) + \sum_i \frac{1\text{TeV}^2}{\Lambda^2} C_i(\mu^2_{\text{EFT}})\tilde{\sigma}_i(\mu^2, \mu^2_{\text{EFT}}) \tag{6.7}$$

where the σ_i are built from the interference of the SM with the EFT amplitudes and

$$\tilde{\sigma}_i(\mu^2, \mu^2_{\text{EFT}}) = \Gamma_{ij}(\mu^2, \mu^2_{\text{EFT}})\sigma_j(\mu^2) \tag{6.8}$$

We will also consider the distributions of the Higgs boson transverse momentum and rapidity, which can be parametrized in a similar fashion.

6.2 Theoretical uncertainties

Considering NLO corrections to SMEFT observables provides a way to reliably estimate the size of missing radiative corrections by considering the dependence of our result on the renormalization and factorization scale. As our predictions are expressed in a double expansion in α_s and in $1/\Lambda$, there are two main classes of uncertainties associated to truncating these expansions. We consider three sources of uncertainties related to the missing orders in the perturbative QCD expansion

- Missing orders in the calculation of the cross section. These are estimated by varying the renormalization and factorization scale by a factor of 2 around the central scale.
- Missing orders in the running of the Wilson coefficients. We estimate this uncertainty by varying μ_{EFT} around the central value by a factor of 2. By separating this running uncertainty from the missing orders the cross section itself, we ensure that no cancellations happen in the global variation. This is especially relevant if we consider global fits, which will use measurements at different scales: the RGE evolution will constitute a major source of uncertainty, which must not be underevaluated.

- PDF uncertainties, which we evaluate using the the procedures provided with our PDF set of choice.

The evaluation of uncertainties associated to the missing orders in the SMEFT expansion are notoriously more difficult to evaluate reliably. A number of discussions and studies have been produced on this topic [108, 193–195]. Let us write down the SMEFT expansion of the cross section up to $1/\Lambda^4$:

$$\sigma = \sigma_{\text{SM}} + \frac{1\text{ TeV}}{\Lambda^2}\sum_i C_i^{(6)}\sigma_i + \frac{(1\text{ TeV})^2}{\Lambda^4}\left[\sum_{ij} C_i^{(6)}C_j^{(6)}\left(\sigma_{i|j}+\sigma_{|ij}\right) + \sum_k C_k^{(8)}\mathfrak{s}_k\right], \tag{6.9}$$

where we restored the dimension label of the Wilson coefficients, the $\sigma_{i|j}$ are the dimension-8 contribution coming from the square of the EFT amplitudes with one dimension-6 operator insertion, which we have calculated, the $\sigma_{|ij}$ come from the interference of the SM amplitude with the SMEFT diagrams with two insertions of dimension six operators and the $\mathfrak{s}_k$ are the interference of the SM amplitude with the SMEFT diagrams with one insertion of a dimension eight operator. Using our calculation, we can explicitly evaluate the $\sigma_{i|j}$ but not the two other contributions. It has been shown using explicit examples that there can be no general arguments based on the $\sigma_{i|j}$ to estimate the size of the other contributions of the same order [194] as this is a model dependent quantity.

6.3 Numerical results in the top sector

6.3.1 Total cross sections

Here we present numerical results for gluon fusion at the LHC at 13 TeV for σ_{SM} and the σ_i. The numbers are quoted with three uncertainties corresponding to the renormalization and factorization scale dependence, the EFT scale dependence, and the PDF uncertainties in that order. Our results are shown in Table 6.1

13 TeV	LO (pb)	NLO (pb)	K-factor
σ_1	$-2.87^{+34.0\%+1.5\%}_{-25.0\%-1.5\%}$	$-4.615^{+24.8\%+1.9\%}_{-20.0\%-1.6\%}$	1.61
σ_2	$2630^{+34.0\%+1.5\%}_{-25.0\%-1.5\%}$	$4080^{+23.9\%+1.9\%}_{-19.6\%-1.6\%}$	1.55
σ_3	$50.2^{+34.0\%+1.5\%}_{-25.0\%-1.5\%}$	$83.0^{+26.0\%+1.9\%}_{-20.6\%-1.6\%}$	1.65

Table 6.1: Results for the total rates at LO and at NLO. The two quoted uncertainties are the scale and PDF uncertainties

We see that the contributions from effective operators have K-factors that are very similar, with a residual scale dependence that is almost identical to the SM. In the following we present an argument which explains this observation. We can describe the total cross section for Higgs boson production to a good accuracy by taking the limit of an infinitely heavy top quark, because most of the production happens near threshold. In this effective theory where the top quark is integrated out, all contributions from SMEFT operators can be described by the same contact interaction $\kappa G^a_{\mu\nu}G_a^{\mu\nu}H$. The Wilson

coefficient κ can be written as

$$\kappa = \kappa_0 + \sum C_i \kappa_i \,, \tag{6.10}$$

where κ_0 denotes the SM contribution and κ_i those corresponding to each operator $\mathcal{O}_i$ in the SMEFT. As a result each σ_i is generated by the same Feynman diagrams both at LO and NLO in the infinite top-mass EFT. The effect of radiative corrections is, however, not entirely universal as NLO corrections to the infinite top-mass EFT amplitudes come both from diagrammatic corrections and corrections to the Wilson coefficients κ_i, which can be obtained by matching the SMEFT amplitude to the infinite top mass amplitude, as illustrated in fig. 6.1. Indeed, each κ_i can be expressed in terms of SMEFT parameters as

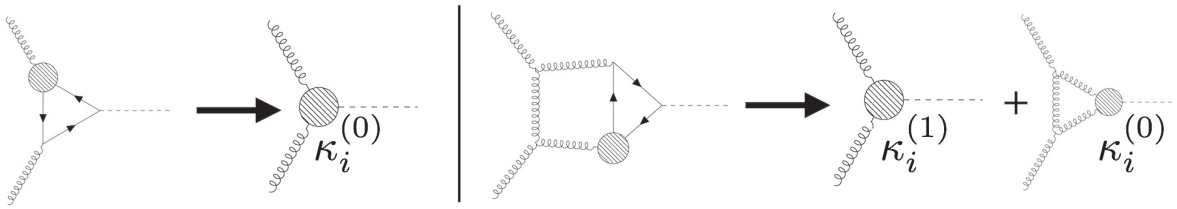


Figure 6.1: Diagrammatic description of the matching between the SMEFT and the infinite top mass EFT at LO (left) and at NLO (right). The NLO amplitude in the infinite top-mass EFT contains two elements: diagrammatic corrections, which contribute universally to the K-factors and Wilson coefficient corrections, which are non-universal.

a perturbative series $\kappa_i = \kappa_i^{(0)} + \alpha_s \kappa_i^{(1)} + \mathcal{O}(\alpha_s^2)$. In the infinite top mass EFT, each K-factor K_i can be decomposed as

$$K_i = K_U + \alpha_s \frac{\kappa_i^{(1)}}{\kappa_i^{(0)}}, \tag{6.11}$$

where K_U is the universal part of the K-factor, which is exactly equal to K_2. By subtracting K_2 to each K_i in the infinite top mass limit numerically (setting $m_t = 10\text{TeV}$), we could extract the ratios $\alpha_s \frac{\kappa_i^{(1)}}{\kappa_i}$ and check explicitly that these non-universal corrections are subdominant compared to the universal diagrammatic corrections, which explains the similarity of the effects of radiative corrections for each contribution.

As could be expected, there is a large hierarchy between the contributions of O_1 and O_3, whose LO contributions are loop induced, and that of O_2 which appears at tree level.[1]

6.3.2 Distributions

We have obtained differential distribution at LO and NLO for the transverse momentum and the rapidity of the Higgs boson in `MadGraph5_aMC@NLO`. Here we present the distributions obtained by matching our NLO prediction to a parton shower to generate events that could be used for detailed simulations of experimental conditions. We parametrize the SMEFT distribution for the variable X truncated at $\mathcal{O}\big(1/\Lambda^2\big)$ in same way as the total cross section:

$$\frac{d\sigma}{dX} = \frac{d\sigma_{\text{SM}}}{dX} + \frac{(1\text{ TeV})^2}{\Lambda^2}\left(\sum_i C_i \frac{d\sigma_i}{dX}\right) \tag{6.12}$$

[1] Indeed $\frac{\sigma_2}{16\pi^2} \simeq \frac{\sigma_1 + \sigma_3}{2}$.

Results for the Higgs transverse momentum (p_T) and rapidity (y) spectra are shown in Figure 6.2. We show the normalized distributions $\frac{1}{\sigma_i}\frac{d\sigma_i}{dX}$ and include an insert showing the ratio of each distribution to the SM distribution. The overall scaling can be obtained from Tables 6.1.

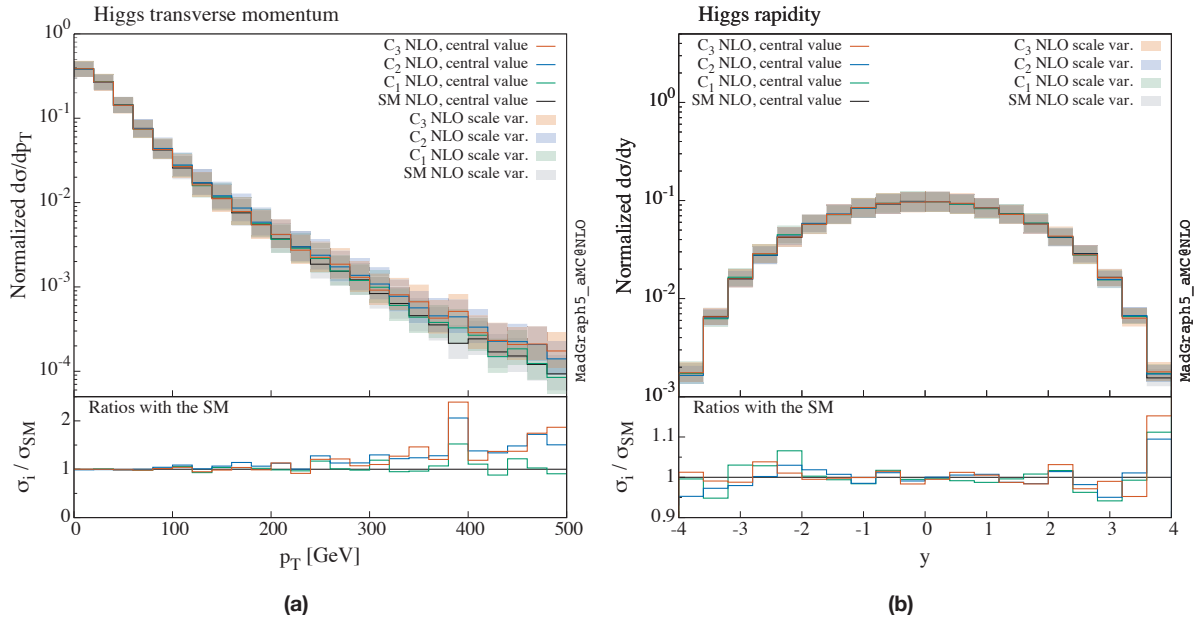


Figure 6.2: Normalized shapes of the Higgs transverse momentum and rapidity distributions predicted by σ_{SM} and σ_i at $13\ \text{TeV}$ including NLO corrections. The bands around the central values show the scale uncertainties of these results (direct sum of the scale dependence and PDF uncertainties). The insert shows the ratio of each SMEFT distribution with the SM distribution.

As we have discussed in the previous section, the transverse momentum spectrum discriminates between the SM production and the effects of the SMEFT operators in the tail of the distribution. Figure 6.2a shows that the internal dynamics of the ggH interaction start to become observable around $250\ \text{GeV}$. On the other hand, Figure 6.2b shows that there is little hope to use the rapidity distribution of the Higgs boson to probe the effects of SMEFT operators. The differences in shape that appear for very forward Higgs bosons, around $|y| > 3.5$, are not statistically significant.

The transverse momentum spectrum of the Higgs boson is only generated by the parton shower at LO. It is therefore expected that large differences should appear at large transverse momentum when including NLO corrections. As shown in Figure 6.3, this is indeed the case for each component of the cross section at transverse momenta above $50\ \text{GeV}$.

While the rapidity distribution is of little phenomenological interest to disentangle possible SMEFT effects, it is still interesting to observe the effect of the NLO corrections on this distribution since it is already fully populated at LO even without matching the calculation to a parton shower, contrary to the transverse momentum spectrum. This makes the rapidity differential K-factor a real measure of an NLO correction, while our description of the p_T spectrum really should be considered a LO prediction since it is the first order where it is nontrivial at the parton level. As a result, as shown in Figure 6.4, the shape correction in the rapidity spectrum is much milder than for the transverse momentum spectrum,

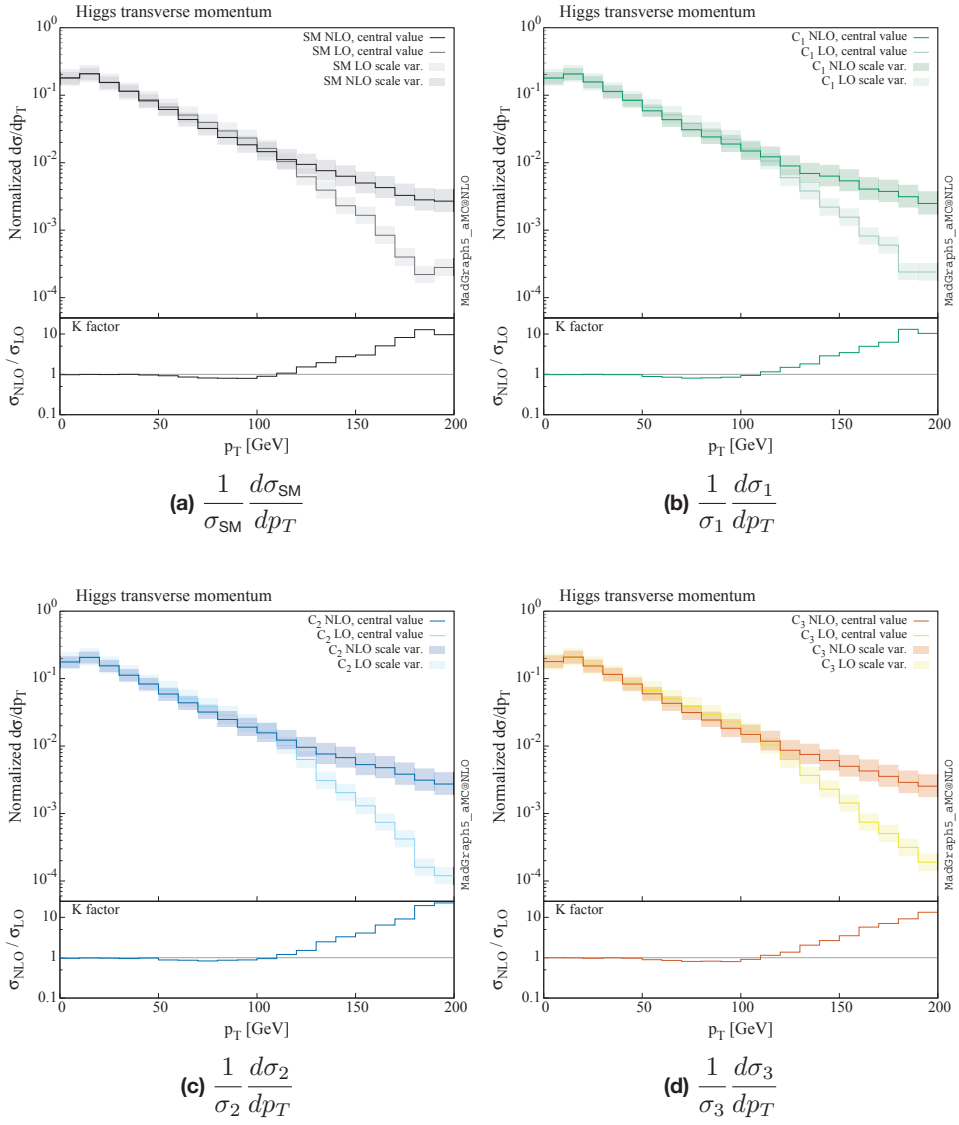


Figure 6.3: Comparison of the LO and NLO shapes of the transverse momentum spectra of σ_{SM} and the σ_i. The inserts show the ratios of the differential cross sections at LO and at NLO.

although the NLO rapidity spectra are noticeably more central.

6.4 Discussion of the effects of light quarks

It was natural as a first step to consider the SMEFT operators that minimally modify the gluon fusion process, *i.e.* those that modify the top loop. However, the SMEFT approach is successful due to its universality and we should not make assumptions by excluding operators that seem less natural. We should therefore investigate whether it would be relevant to include the effects of the variants $O_1^{(f)}$ and $O_3^{(f)}$ of operators O_1 and O_3 for other quark flavors.

The constraints on the top-quark O_1 and O_3 have been shown to be set in a complementary way by gluon fusion and $\bar{t}tH$ associated production in [118]. This statement would however not carry on to

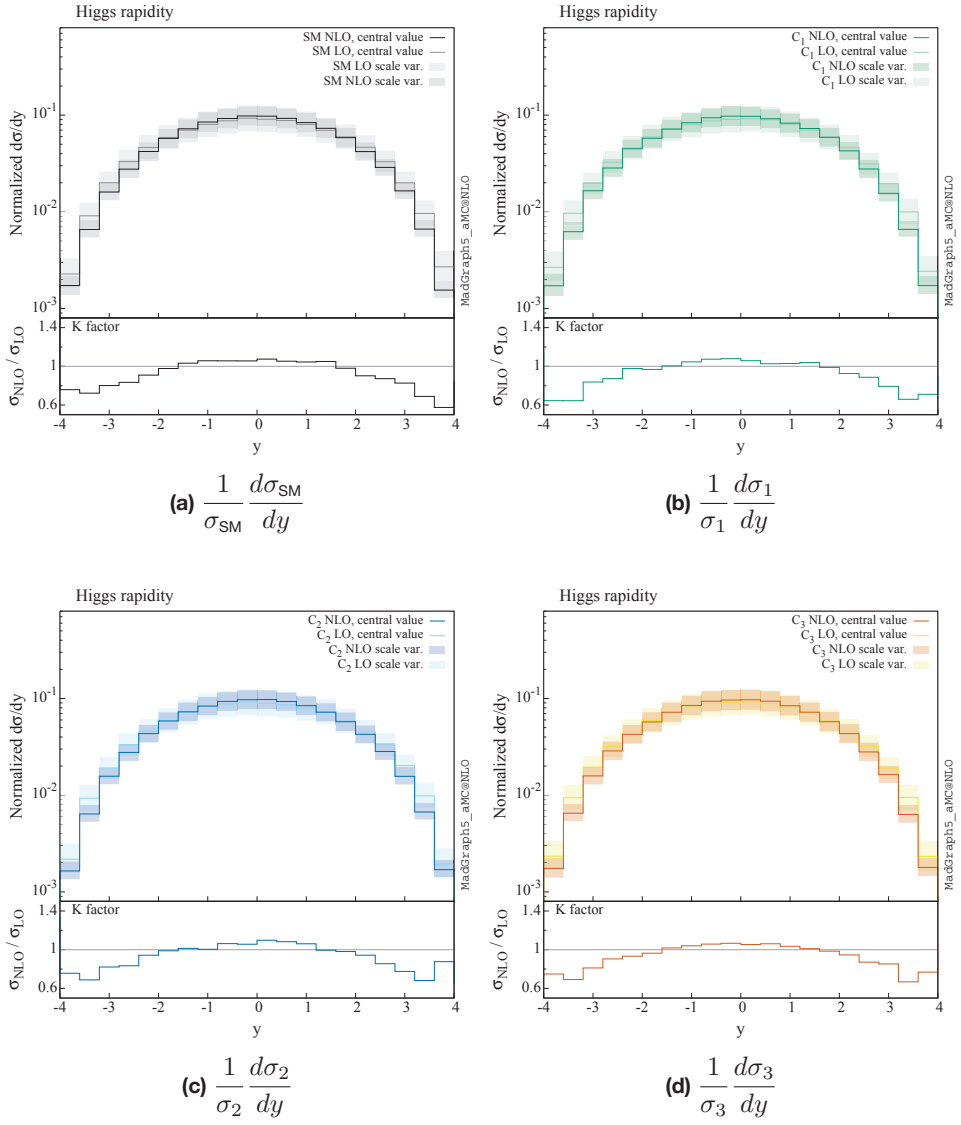


Figure 6.4: Comparison of the LO and NLO shapes of the transverse momentum spectra of σ_{SM} and the σ_i. The inserts show the ratios of the differential cross sections at LO and at NLO.

the lighter flavors and we can make the argument that any values of the Wilson coefficients of $O_1^{(f)}$ and $O_3^{(f)}$ that would yield an observable effect in gluon fusion through a light quark loop would already have been constrained by their contribution to other production modes. Indeed, if $O_1^{(f)}$ brings a significant contribution to gluon fusion, it would bring a much stronger contribution through the quark-antiquark annihilation channel. Similary, if $O_3^{(f)}$ is relevant for gluon fusion, it also generates several new production modes happening at tree level ($q\bar{q} \to Hg$, $gq \to Hq$, $gg \to Hq\bar{q}$) which would be much more constraining. The contribution of all quark initiated processes would be suppressed in the case of the bottom quark, but both operators would yield a significant contribution to $H\bar{b}b$. A dedicated study of the contribution of the bottom-quark related operators would still be interesting and might warrant their inclusion in our prediction, but we can already safely neglect all other flavors.

7 Conclusion

In this thesis, we have presented two original results that expand the existing theoretical toolbox for the study of the properties of the Higgs boson at the LHC. The motivation for this work is rooted in the challenges of modern particle physics, in which the absence of new physics signals combined with the ever expanding body of BSM models makes a compelling case for precision measurements as a way to study the SM and to constrain BSM effects in the most model-independent way possible. As theorists, playing our part in this enterprise today implies tackling multi-loop calculations to make precise predictions. To this end, we used modern mathematical techniques that are interesting in their own right but have also very pragmatic applications for the world of particle physics. Using these tools, we have been able to provide results that we hope will help shed more light on the Higgs sector.

The first result presented in this thesis was the extraction of the power-suppressed correction to the bottom quark Yukawa in the HEFT, which corresponds to the infinite top mass limit of the SM. This correction is needed to perform the complete calculation of the $\mathcal{O}(\alpha_s^4 y_t y_b)$ correction to the production cross section of a Higgs boson in association with two bottom quarks in the HEFT, which, while subleading, is expected to reduce the large theoretical uncertainties in the state-of-the-art calculation. To this end, we calculated the $\mathcal{O}(\alpha_s^2 y_t)$ contribution to a form factor contributing to the decay of a Higgs boson to a pair of bottom quarks in the SM and in the HEFT, which allowed us to match the two theories and extract the desired correction. The HEFT calculation was the combination of one-loop diagrams and of a tree-level diagram which was proportional to the HEFT bottom quark Yukawa and we performed the calculation by reducing the one-loop diagram to master integrals which we evaluated exactly using the Feynman parametrization. One of the masters was a triangle integral which is expressed in terms of MPLs and we used techniques exploiting their Hopf algebra structure to perform the integration and obtain the HEFT form factor. We then turned to the SM form factor, which is the sum of two two-loop diagrams, whose Feynman integrals we also reduced as linear combinations of a

basis of two-loop master integrals. We then evaluated these integrals in an expansion in inverse powers of the top quark mass using the expansion by regions, using simple tricks for the majority of integrals and exploiting Hopf-algebra-based techniques again for the most complicated cases. This calculation showed in particular that these complicated integrals are in fact proportional to the triangle integral that appeared in the HEFT in the infinite top mass limit. Using the values for our master integrals, we could obtain the form factor in the SM and compare it to the value obtained in the HEFT. This showed indeed that the two form factors have the same structure up to power-suppressed terms after renormalization and we could determine the value of the power-suppressed correction to the bottom quark Yukawa in the HEFT in terms of SM parameters. This calculation was of course a means to an end and we now plan to use it to improve the theoretical prediction for $Hb\bar{b}$ associated production in the SM.

A second set of results presented in this thesis correspond to our study of the gluon fusion process at NLO in the SMEFT. We evaluated the two-loop Feynman diagrams that correspond to the virtual corrections to this process in terms of seventeen master two loop integrals. We used the expressions of these integrals provided in the literature to obtain the expression of the two-loop bare amplitude and exploited the fact that its one-loop UV divergences as well as the infrared divergences could readily be determined to fix the value of a two-loop counterterm that mixes the gluon fusion operator and the chromomagnetic operator. Using this result, we could express the renormalized amplitude and use its IR subtracted combined with automatically generated real emission corrections in the `MadGraph5_aMC@NLO` framework to numerically evaluate cross sections and kinematic distributions. We presented the results of this numerical calculation and discussed the effects of the NLO corrections on physical observables. In particular, we highlighted the universality of these corrections for each piece of the total cross section and proposed an explanation for this observation based on the infinite top mass limit of the amplitude. As this region of phase space is phenomenologically relevant if we want to disentangle the effects of each operator, this shows that the exact evaluation of the NLO corrections was indeed important. These results provide tools to constrain new physics in the Higgs sector using total and differential cross section measurements.

As the LHC is only at the beginning of its data taking period, we have exciting times ahead in which we will put the elementary interactions and particles under unprecedented scrutiny and probe the details of electroweak symmetry breaking. There is much we can do as theorists to be part of this endeavor as we further expand our work at the precision frontier, which also provides an amazing opportunity to better understand the perturbative structure of the quantum field theories that are the basis of our understanding of the microscopic world.

A

Shortcomings of Δ_Φ

Let us give an example of why the coproduct Δ_Φ defined in Section 3.3.2 is not well defined beyond the simplest case where we consider a single variable and only constant entries. Let us consider the MPL $G(\frac{1}{x}, 1; x)$. We can use the full technology provided by the Goncharov coproduct to put this function into canonical form:

$$G\left(\frac{1}{x}, 1; x\right) = G(-1, 1; x) - G(0, -1; x) - G(0, 1; x) + G(1, -1; x) + G(1, 1; x). \tag{A.1}$$

Let us now evaluate the coproduct Δ_Φ on each side of the equation:

$$\begin{aligned}\Delta_\Phi\left(G\left(\frac{1}{x}, 1; x\right)\right) = {} & 1 \otimes G(-1, 1; x) - 1 \otimes G(0, -1; x) - 1 \otimes G(0, 1; x) + 1 \otimes G(1, -1; x) \\ & + 1 \otimes G(1, 1; x) + G(-1; x) \otimes G(1; x) + G(1; x) \otimes G(1; x) + G(-1, 1; x) \otimes 1 \\ & - G(0, -1; x) \otimes 1 - G(0, 1; x) \otimes 1 + G(1, -1; x) \otimes 1 + G(1, 1; x) \otimes 1,\end{aligned} \tag{A.2}$$

while

$$\begin{aligned}\Delta_\Phi\Big(G(-1, 1; x) - G(0, -1; x) - G(0, 1; x) + G(1, -1; x) + G(1, 1; x)\Big) = {} & 1 \otimes G(-1, 1; x) \\ & - 1 \otimes G(0, -1; x) - 1 \otimes G(0, 1; x) + 1 \otimes G(1, -1; x) + 1 \otimes G(1, 1; x) \\ & + G(-1; x) \otimes G(1; x) + G(1; x) \otimes G(-1; x) + G(1; x) \otimes G(1; x) + G(-1, 1; x) \otimes 1 \\ & - G(0, -1; x) \otimes 1 - G(0, 1; x) \otimes 1 + G(1, -1; x) \otimes 1 + G(1, 1; x) \otimes 1,\end{aligned} \tag{A.3}$$

which is to say

$$\Delta_\Phi\left(G\left(\frac{1}{x},1;x\right)\right)-\Delta_\Phi\left(G(-1,1;x)-G(0,-1;x)-G(0,1;x)+G(1,-1;x)+G(1,1;x)\right)$$
$$=-G(1;x)\otimes G(-1;x). \tag{A.4}$$

B

Details of code implementation

B.1 Generating diagrams with QGRAF and qgraf-xml-parser

One of the most popular tools available for Feynman diagram generation is the program QGRAF [125]. This code lets the user specify Feynman rules as diagrammatic rules and builds the diagrams as an abstract representation of the diagram. It is then up to the user to specify how to output this representation and to translate it to analytic expressions. To this end, we designed an output format for QGRAF that encodes feynman diagrams as XML objects describing the vertices and the propagators. These XML files were then read by our Python program qgraf-xml-parser to translate the diagrams into FORM expressions. The XML format was chosen because it is both human-readable, meaning that we could look at it and understand the diagrams described in the code, and the machine-readable, meaning that the files could be imported into Python directly as structured objects, allowing for efficient processing. Each diagram is imported into qgraf-xml-parser as an object containing a list of vertices and a list of propagator, where each propagator contains information about the vertices it is connected to and vice versa. The program contains FORM [127] expressions for the Feynman rules of each vertex and propagator in the form of strings and builds the expression for the diagram by first building the expression for fermion lines, following them vertex-by-vertex, and then multiplying all the chains of Dirac matrices by the bosonic vertices and propagators. The output FORM file can then be processed by FORM, which can very efficiently calculate all the index contractions in the amplitude as well as calculate the Dirac matrix traces. The FORM output was formatted to correspond to Mathematica syntax and was imported in this program for the further steps of processing.

More anecdotically, as qgraf-xml-parser already had all the necessary structure to represent Feynman diagrams abstractly in Python, we developed a small spinoff called qgraf-xml-drawer [196], which automatically writes (Lua)LaTeX files to draw Feynman diagrams specified in our XML format using the package TikZ-Feynman [197].

B.2 Family reduction in practice: SAIF

The integrals $I(n_1, \ldots, n_N)$ defined above for a given family provide us with very useful tools for calculating loop integrals as we shall see in the next section. We therefore developped the `Mathematica` package `SAIF` (**S**orting **A**mplitudes **I**nto **F**amilies) to assign a topology to our diagrams and express all integrals as linear combinations of $I(n_1, \ldots, n_N)$. There are two main functions in this package, `FindTopology`, and `ToTopology`. The first function takes as argument the expression for the integrand of a loop amplitude and a list of possible topologies that need to be determined by hand. Topologies are defined as lists of propagators along with kinematic information on the external legs. The package contains helper functions to find the replacement rules to express numerators as linear combinations of denominators, as well as to generate all the possible permutations of identical legs for a given topology. The `FindTopology` function works by extracting the denominators from a Feynman diagram and trying to match them to the list of topologies (including all permutations of legs) by looping over them. Of course, the choice of loop momenta is not unique so for each topology due to shift invariance in dimensional regularization, so for each topology, we consider all shifts of our loop momenta k_i of the form

$$k_i \to A_i^j k_j + c_k p_k, \tag{B.1}$$

where the p_k are the external momenta and A_i^j, c_k are taken from $\{0, \pm 1\}$ with $|\det(A)| = 1$. This is sufficient to cover all momentum shifts arising in practice. Once a match is found, the function outputs the corresponding topology as well as the required loop momentum shift to put the diagram propagators into the same form as the family. The function `ToTopology` takes as input a Feynman diagram and the output of `FindTopology`. It applies the loop momentum shift to the diagram, then the numerator reduction rules, which transforms the diagram into a linear combination of integrals with only denominators. It is then only a matter of matching each term in the sum as an $\alpha D_1^{n_1} \ldots D_N^{n_N}$ to express the diagram as a linear combination of $I(n_1, \ldots, n_N)$.

C

Two-loop master integrals in the $H \to b\bar{b}$ calculation

In this appendix, we show the results for the 23 master integrals that appear in the calculation of the two-loop form factor expanded in the infinite top mass limit that we evluated in Section 4.2.3. The integrals showed here are defined with a $\overline{\text{MS}}$ prefactor

$$J(n_1,\ldots,n_7) = (4\pi)^{-2\epsilon} e^{2\epsilon\gamma} \int \frac{d^d k_1}{(2\pi)^d} \frac{d^d k_2}{(2\pi)^d} \frac{1}{D_1^{n_1}\ldots D_7^{n_7}}, \tag{C.1}$$

where

$$\begin{aligned} &D_1 = k_2^2 - m_b^2, && D_2 = (k_2 - p_1)^2, && D_3 = (k_2 + p_2)^2, \\ &D_4 = (k_1 - p_1)^2 - m_t^2, && D_5 = (k_1 + p_2)^2 - m_t^2, && D_6 = (k_1 - k_2)^2 - m_t^2, \\ & && D_7 = (k_1 + k_2)^2 - m_t^2. \end{aligned} \tag{C.2}$$

We express our integrals in terms of $t = m_b^2/m_t^2$

$$J(0,0,0,1,0,1,0) = \left(-\frac{1}{256\pi^4}\right)\frac{1}{\epsilon^2} + \left(-\frac{1}{128\pi^4}\right)\frac{1}{\epsilon} - \frac{1}{1536\pi^2} - \frac{3}{256\pi^4}$$

$$J(1,0,0,1,0,0,0) = \frac{1}{\epsilon^2}\left(t\left(-\frac{1}{256\pi^4}\right)\right) + \frac{1}{\epsilon}\left(t\frac{\log(t)-2}{256\pi^4}\right) + t\left(-\frac{3\log^2(t)-12\log(t)+\pi^2+18}{1536\pi^4}\right)$$

$$J(0,0,1,1,0,1,0) = \left(-\frac{1}{256\pi^4}\right)\frac{1}{\epsilon^2} + \frac{1}{\epsilon}\left(t\left(-\frac{1}{1024\pi^4 r(r+1)}\right) - \frac{3}{256\pi^4}\right)$$
$$+ \frac{1}{9216\pi^4 r^2(r+1)^2}t^2 + \frac{1}{2048\pi^4 r(r+1)}t - \frac{1}{1536\pi^2} - \frac{7}{256\pi^4}$$

$$J(0,0,2,1,0,1,0) = \frac{1}{512\pi^4}\frac{1}{\epsilon^2} + \frac{1}{\epsilon}\left(t^2\frac{1}{15360\pi^4 r^2(r+1)^2} + t\left(-\frac{1}{1536\pi^4 r(r+1)}\right) - \frac{1}{512\pi^4}\right)$$
$$+ \frac{17}{115200\pi^4 r^2(r+1)^2}t^2 + \frac{1}{9216\pi^4 r(r+1)}t + \frac{18+\pi^2}{3072\pi^4}$$

$$J(0,0,1,2,0,1,0) = \left(-\frac{1}{512\pi^4}\right)\frac{1}{\epsilon^2} + \left(-\frac{1}{512\pi^4}\right)\frac{1}{\epsilon} + \left(-\frac{1}{18432\pi^4 r^2(r+1)^2}\right)t^2$$
$$+ \frac{1}{1024\pi^4 r(r+1)}t + \frac{-6-\pi^2}{3072\pi^4}$$

$$J(0,0,0,1,1,1,0) = \left(-\frac{1}{256\pi^4}\right)\frac{1}{\epsilon^2} + \frac{1}{\epsilon}\left(t^2\left(-\frac{1}{15360\pi^4 r^2(r+1)^2}\right) + t\frac{1}{1536\pi^4 r(r+1)} - \frac{1}{256\pi^4}\right)$$
$$+ \left(-\frac{1}{7680\pi^4 r^2(r+1)^2}\right)t^2 + \frac{1}{1536\pi^4 r(r+1)}t + \frac{-6-\pi^2}{1536\pi^4}$$

$$J(0,1,1,1,0,0,0) = \left(-\frac{1}{256\pi^4}\right)\frac{1}{\epsilon^2} + \frac{1}{\epsilon}\frac{-\log(r(r+1)) + \log(t) - 3}{256\pi^4}$$
$$+ \frac{2\log(r(r+1))\log(t) - \log^2(r(r+1)) - 6\log(r(r+1)) - \log^2(t) + 6\log(t) - 14}{512\pi^4}$$

$$J(2,0,0,0,1,1,0) = \left(-\frac{1}{512\pi^4}\right)\frac{1}{\epsilon^2} + \frac{1}{\epsilon}\frac{2\log(t)-1}{512\pi^4} + t^2\frac{6\log(t)-5}{9216\pi^4}$$
$$+ \frac{-6\log^2(t) - \pi^2 + 18}{3072\pi^4} + t\frac{2\log(t)-3}{1024\pi^4}$$

$$J(1,0,0,1,0,1,0) = \frac{1}{\epsilon^2}\left(t\left(-\frac{1}{512\pi^4}\right) - \frac{1}{256\pi^4}\right) + \frac{1}{\epsilon}\left(t^3\left(-\frac{1}{3072\pi^4}\right) + t\frac{4\log(t)-5}{1024\pi^4} - \frac{3}{256\pi^4}\right)$$
$$+ t^3\frac{10\log(t)-21}{30720\pi^4} + t^2\frac{6\log(t)-11}{4608\pi^4} + t\frac{-12\log^2(t)+24\log(t)-2\pi^2+9}{6144\pi^4} + \frac{-42-\pi^2}{1536\pi^4}$$

$$J(2,0,0,1,0,1,0) = \left(-\frac{1}{512\pi^4}\right)\frac{1}{\epsilon^2} + \frac{1}{\epsilon}\frac{2\log(t)-1}{512\pi^4} + t^2\frac{6\log(t)-5}{9216\pi^4}$$
$$+ \frac{-6\log^2(t) - \pi^2 + 18}{3072\pi^4} + t\frac{2\log(t)-3}{1024\pi^4}$$

$$J(1,0,0,1,1,0,0) = \frac{1}{\epsilon^2}\left(t\left(-\frac{1}{256\pi^4}\right)\right) + \frac{1}{\epsilon}\left(t^3\left(-\frac{1}{15360\pi^4 r^2(r+1)^2}\right) + t^2\frac{1}{1536\pi^4 r(r+1)} + t\frac{\log(t)-1}{256\pi^4}\right)$$
$$+ t^3\frac{\log(t)-2}{15360\pi^4 r^2(r+1)^2} + t^2\left(-\frac{\log(t)-1}{1536\pi^4 r(r+1)}\right) + t\left(-\frac{3\log^2(t)-6\log(t)+\pi^2+6}{1536\pi^4}\right)$$

$$J(0,1,1,1,1,0,0) = \left(-\frac{1}{256\pi^4}\right)\frac{1}{\epsilon^2} + \frac{1}{\epsilon}\left(\frac{-\log(r(r+1)) + \log(t) - 2}{256\pi^4}\right.$$
$$+t^2\left(-\frac{1}{15360\pi^4 r^2(r+1)^2}\right) + t\frac{1}{1536\pi^4 r(r+1)}\Big) + t^2\frac{-\log(r(r+1)) + \log(t) - 3}{15360\pi^4 r^2(r+1)^2}$$
$$+\frac{2\log(r(r+1))\log(t) - \log^2(r(r+1)) - 4\log(r(r+1)) - \log^2(t) + 4\log(t) - 8}{512\pi^4}$$
$$+t\frac{\log(r(r+1)) - \log(t) + 2}{1536\pi^4 r(r+1)}$$

$$J(1,0,0,1,1,1,0) = \left(-\frac{1}{512\pi^4}\right)\frac{1}{\epsilon^2} + \frac{1}{\epsilon}\left(t^2\left(-\frac{1}{15360\pi^4 r^2(r+1)^2}\right) + t\frac{1}{1536\pi^4 r(r+1)}\right.$$
$$+t^3\frac{30r^4 + 60r^3 + 23r^2 - 7r + 2}{92160\pi^4 r^2(r+1)^2} - \frac{1}{512\pi^4}\Big) + t\frac{18r^2 - 12(r+1)\log(t)r + 18r - 1}{6144\pi^4 r(r+1)}$$
$$+t^3\frac{78r^4 + 156r^3 + 61r^2 - 17r - \left(30r^4 + 60r^3 + 23r^2 - 7r + 2\right)\log(t) + 5}{92160\pi^4 r^2(r+1)^2}$$
$$+t^2\frac{300r^4 + 600r^3 + 215r^2 - 90\left(4r^3 + 8r^2 + 3r - 1\right)\log(t)r - 85r - 43}{552960\pi^4 r^2(r+1)^2} - \frac{1}{3072\pi^2} - \frac{1}{512\pi^4}$$

$$J(2,0,0,1,1,1,0) = t\left(-\frac{\left(6r^2 + 6r - 1\right)\left(2\log(t) - 1\right)}{12288\pi^4 r(r+1)}\right)$$
$$+t^2\left(-\frac{\left(30r^4 + 60r^3 + 25r^2 - 5r + 1\right)\left(3\log(t) - 1\right)}{138240\pi^4 r^2(r+1)^2}\right) - \frac{\log(t)}{512\pi^4} + \frac{1}{512\pi^4}$$

$$J(1,0,0,1,1,2,0) = \frac{1}{\epsilon}\left(t^3\left(-\frac{90r^4 + 180r^3 + 77r^2 - 13r + 2}{215040\pi^4 r^2(r+1)^2}\right)\right) + t\left(-\frac{2r^2 - 12(r+1)\log(t)r + 2r + 5}{18432\pi^4 r(r+1)}\right)$$
$$+t^3\left(-\frac{792r^4 + 1584r^3 + 686r^2 - 106r - 3\left(90r^4 + 180r^3 + 77r^2 - 13r + 2\right)\log(t) + 17}{645120\pi^4 r^2(r+1)^2}\right)$$
$$+t^2\left(-\frac{168r^4 + 336r^3 + 172r^2 - 90\left(8r^3 + 16r^2 + 7r - 1\right)\log(t)r + 4r - 47}{1382400\pi^4 r^2(r+1)^2}\right) + \frac{1}{256\pi^4}$$

$$J(1,0,0,2,1,1,0) = \frac{1}{\epsilon}\left(t^2\frac{1}{15360\pi^4 r^2(r+1)^2} + t\left(-\frac{1}{3072\pi^4 r(r+1)}\right)\right.$$
$$+t^3\left(-\frac{10r^4 + 20r^3 + 7r^2 - 3r + 1}{35840\pi^4 r^2(r+1)^2}\right) + \frac{1}{512\pi^4}\Big) + t\frac{3r(r+1)\log(t) - 2\left(r^2 + r + 1\right)}{4608\pi^4 r(r+1)}$$
$$+t^3\left(-\frac{176r^4 + 352r^3 + 126r^2 - 50r - 6\left(10r^4 + 20r^3 + 7r^2 - 3r + 1\right)\log(t) + 17}{215040\pi^4 r^2(r+1)^2}\right)$$
$$+t^2\frac{-36r^4 - 72r^3 - 19r^2 + 30\left(3r^3 + 6r^2 + 2r - 1\right)\log(t)r + 17r + 29}{230400\pi^4 r^2(r+1)^2}$$

$$J(1,0,1,1,0,1,0) = \left(-\frac{1}{512\pi^4}\right)\frac{1}{\epsilon^2} + \frac{1}{\epsilon}\left(\frac{\log(t)}{256\pi^4} - \frac{5}{512\pi^4}\right) + t\frac{-26r^2 - 26r + 6\left(2r^2 + 2r - 1\right)\log(t) + 19}{18432\pi^4 r(r+1)}$$
$$+t^2\left(-\frac{37r^4 + 74r^3 + 23r^2 - 14r - 5\left(6r^4 + 12r^3 + 4r^2 - 2r + 1\right)\log(t) + 12}{230400\pi^4 r^2(r+1)^2}\right)$$
$$-\frac{\log^2(t)}{512\pi^4} + \frac{\log(t)}{128\pi^4} - \frac{1}{3072\pi^2} - \frac{5}{512\pi^4}$$

$$J(2,0,1,1,0,1,0) = \frac{1}{\epsilon^2}\left(\frac{1}{t}\left(-\frac{1}{512\pi^4}\right)\right)$$
$$+\frac{1}{\epsilon}\left(t\left(-\frac{1}{30720\pi^4 r^2(r+1)^2}\right)+t^2\frac{10r^4+20r^3+7r^2-3r+1}{107520\pi^4 r^2(r+1)^2}+\frac{1}{t}\frac{\log(t)}{512\pi^4}+\frac{1}{3072\pi^4 r(r+1)}\right)$$
$$+\frac{3110r^4+6220r^3+1967r^2-1143r+486}{22579200\pi^4 r^2(r+1)^2}t^2$$
$$+t\left(-\frac{81r^4+162r^3+49r^2-32r-15\left(6r^4+12r^3+4r^2-2r+1\right)\log(t)+60}{460800\pi^4 r^2(r+1)^2}\right)$$
$$-\frac{\log(t)}{3072\pi^4 r(r+1)}+\frac{1}{1536\pi^4 r(r+1)}+\frac{1}{t}\left(-\frac{3\log^2(t)+\pi^2}{3072\pi^4}\right)\frac{\log(t)}{1536\pi^4}-\frac{5}{4608\pi^4}$$
$$J(1,0,1,2,0,1,0) = t\frac{14r^2+14r-6\left(2r^2+2r-1\right)\log(t)-13}{36864\pi^4 r(r+1)}$$
$$+t^2\frac{44r^4+88r^3+26r^2-18r-10\left(6r^4+12r^3+4r^2-2r+1\right)\log(t)+19}{460800\pi^4 r^2(r+1)^2}$$
$$-\frac{\log(t)}{512\pi^4}+\frac{3}{512\pi^4}$$
$$J(1,1,1,1,0,0,0) = \frac{1}{\epsilon}\left(\frac{1}{t}\frac{r(r+1)\left(-3G(-1,-1,r)-3G(-1,0,r)+3G(0,-1,r)+3G(0,0,r)+2\pi^2\right)}{768\pi^4(2r+1)}\right)$$
$$+\frac{1}{t}\frac{1}{1536\pi^4(2r+1)}\left(r(r+1)\left(\pi^2 G(-1,r)+8\pi^2 G\left(-\frac{1}{2},r\right)-\pi^2 G(0,r)-6G(-1,-1,r)-6G(-1,0,r)\right.\right.$$
$$+6G(0,-1,r)+6G(0,0,r)-6G(-1,-1,-1,r)-6G(-1,-1,0,r)-6G(-1,0,-1,r)-6G(-1,0,0,r)$$
$$-12G\left(-\frac{1}{2},-1,-1,r\right)-12G\left(-\frac{1}{2},-1,0,r\right)+12G\left(-\frac{1}{2},0,-1,r\right)+12G\left(-\frac{1}{2},0,0,r\right)$$
$$+6G(0,-1,-1,r)+6G(0,-1,0,r)+6G(0,0,-1,r)+6G(0,0,0,r)+6G(-1,-1,r)\log(t)$$
$$\left.\left.+6G(-1,0,r)\log(t)-6G(0,-1,r)\log(t)-6G(0,0,r)\log(t)-4\pi^2\log(t)+30\zeta(3)+4\pi^2\right)\right)$$
$$J(0,1,1,1,1,1,0) = t^2\frac{\log(r(r+1))-\log(t)+1}{46080\pi^4 r^2(r+1)^2}+t\frac{-2\log(r(r+1))+2\log(t)-3}{12288\pi^4 r(r+1)}$$
$$+\frac{\log(r(r+1))}{512\pi^4}-\frac{\log(t)}{512\pi^4}+\frac{3}{512\pi^4}$$

$$J(1,1,1,1,1,0,0) = \frac{1}{\epsilon}\left(\frac{1}{t}\frac{r(r+1)\left(-3G(-1,-1,r)-3G(-1,0,r)+3G(0,-1,r)+3G(0,0,r)+2\pi^2\right)}{768\pi^4(2r+1)}\right)$$
$$+\frac{1}{t}\frac{1}{1536\pi^4(2r+1)}\left(r(r+1)\left(\pi^2 G(-1,r)+8\pi^2 G\left(-\frac{1}{2},r\right)-\pi^2 G(0,r)\right.\right.$$
$$-6G(-1,-1,-1,r)-6G(-1,-1,0,r)-6G(-1,0,-1,r)-6G(-1,0,0,r)-12G\left(-\frac{1}{2},-1,-1,r\right)$$
$$-12G\left(-\frac{1}{2},-1,0,r\right)+12G\left(-\frac{1}{2},0,-1,r\right)+12G\left(-\frac{1}{2},0,0,r\right)+6G(0,-1,-1,r)$$
$$+6G(0,-1,0,r)+6G(0,0,-1,r)+6G(0,0,0,r)+6G(-1,-1,r)\log(t)$$
$$\left.\left.+6G(-1,0,r)\log(t)-6G(0,-1,r)\log(t)-6G(0,0,r)\log(t)-4\pi^2\log(t)+30\zeta(3)\right)\right)$$
$$+t\frac{-3G(-1,-1,r)-3G(-1,0,r)+3G(0,-1,r)+3G(0,0,r)+2\pi^2}{46080\pi^4 r(r+1)(2r+1)}$$
$$+\frac{rG(-1,-1,r)}{1536\pi^4(r+1)(2r+1)}+\frac{rG(-1,0,r)}{1536\pi^4(r+1)(2r+1)}-\frac{rG(0,-1,r)}{1536\pi^4(r+1)(2r+1)}$$
$$-\frac{rG(0,0,r)}{1536\pi^4(r+1)(2r+1)}+\frac{G(-1,-1,r)}{1536\pi^4(r+1)(2r+1)}$$
$$+\frac{G(-1,0,r)}{1536\pi^4(r+1)(2r+1)}-\frac{G(0,-1,r)}{1536\pi^4(r+1)(2r+1)}-\frac{G(0,0,r)}{1536\pi^4(r+1)(2r+1)}$$
$$-\frac{r}{2304\pi^2(r+1)(2r+1)}-\frac{1}{2304\pi^2(r+1)(2r+1)}$$
$$J(1,1,1,1,1,1,0) = \frac{1}{\epsilon}\left(t\left(-\frac{r^2+r-1}{15360\pi^4 r(r+1)}\right)\right)$$
$$+\frac{1}{t}\left(-\frac{r(r+1)\left(-3G(-1,-1,r)-3G(-1,0,r)+3G(0,-1,r)+3G(0,0,r)+2\pi^2\right)}{1536\pi^4(2r+1)}\right)$$
$$-\frac{G(-1,-1,r)}{6144\pi^4(2r+1)}-\frac{G(-1,0,r)}{6144\pi^4(2r+1)}+\frac{G(0,-1,r)}{6144\pi^4(2r+1)}+\frac{G(0,0,r)}{6144\pi^4(2r+1)}$$
$$+\frac{-86r^2-86r+91}{921600\pi^4 r(r+1)}t-\frac{r\log(t)}{1536\pi^4(2r+1)}-\frac{\log(t)}{3072\pi^4(2r+1)}+\frac{1}{9216\pi^2(2r+1)}+\frac{1}{1152\pi^4}$$

D

Details of the $gg \to H$ calculation

D.1 Definition of the master integrals

Three of the master integrals used in the calculation of the two loop amplitude for gluon fusion in the SMEFT are ambiguous because they contain numerators, which requires choosing an explicit parametrization. Here we provide their definitions:

$$\text{Int}_c = \int \frac{d^d k_1}{(2\pi^d)} \frac{d^d k_2}{(2\pi^d)} \frac{k_1 \cdot k_2}{(k_1^2 - m_t^2) k_2^2 \left((p_1 - k_1)^2 - m_t^2\right) (p_1 - k_1 + k_2)^2 - m_t^2 \left((p_2 + k_1 - k_2)^2 - m_t^2\right)}$$

$$\text{Int}_h = int \frac{d^d k_1}{(2\pi^d)} \frac{d^d k_2}{(2\pi^d)} \frac{k_1 \cdot k_2}{(k_1^2 - m_t^2)(k_2^2 - m_t^2) \left((p_1 - k_1 + k_2)^2 - m_t^2\right) (p_2 + k_1 - k2)^2}$$

$$\text{Int}_k = \frac{d^d k_1}{(2\pi^d)} \frac{d^d k_2}{(2\pi^d)} \frac{k_1 \cdot k_2}{k_2^2 (k_1 - m_t^2) \left((p_1 + p_2 - k_1 - k_2)^2 - m_t^2\right)}$$

D.2 Basis rotation at two loops

In Section 5.2.3, we dicussed the reduction to master integrals of the two-loop scalar integrals that appeared in the calculation of the two-loop diagrams relevant for the calculation of the cross-section of Higgs production through gluon fusion at NLO in the SMEFT. We expressed the set of master integrals that was produced by `LiteRed` to the basis used in the calculation of the SM cross-section for gluon fusion at two loops by using `SAIF` to reduce the known masters in our `LiteRed` basis and inverting the linear system. Here we show the relation between the bases shown in Equations D.1, D.3, and D.5. We

use the variable $\mathfrak{s} = s/mt^2$.

Basis rotation for T_1 All the diagrams whose integrals fall in T_1 can be expressed as a linear combination of 12 master integrals forming the basis B_1:

$$\begin{array}{llll} \mathcal{J}_{T_1}(0,0,0,0,0,1,1), & \mathcal{J}_{T_1}(0,0,0,1,0,1,1), & \mathcal{J}_{T_1}(0,0,0,1,1,1,1), & \mathcal{J}_{T_1}(0,0,1,0,0,1,1), \\ \mathcal{J}_{T_1}(0,0,1,0,1,1,1), & \mathcal{J}_{T_1}(0,0,2,0,0,1,1), & \mathcal{J}_{T_1}(0,1,0,1,0,1,1), & \mathcal{J}_{T_1}(0,1,0,2,0,1,1), \\ \mathcal{J}_{T_1}(0,2,0,1,0,1,1), & \mathcal{J}_{T_1}(1,0,1,0,0,0,1), & \mathcal{J}_{T_1}(1,0,1,1,0,1,0), & \mathcal{J}_{T_1}(1,0,1,1,0,1,1), \end{array} \tag{D.1}$$

which we could express in linear combinations of the following 12 known masters:

$$\begin{array}{llllll} \mathrm{Int}_{\mathrm{bmb}}, & \mathrm{Int}_e, & \mathrm{Int}_f, & \mathrm{Int}_g, & \mathrm{Int}_h, & \mathrm{Int}_i, \\ \mathrm{Int}_j, & \mathrm{Int}_k, & \mathrm{Int}_{\mathrm{tb}}, & \mathrm{Int}_{\mathrm{tmb}}, & \mathrm{Int}_{\mathrm{tt}}, & \mathrm{Int}_{\mathrm{ttri}}. \end{array} \tag{D.2}$$

Basis for T_2 All the diagrams whose integrals fall in T_2 can be expressed as a linear combination of 14 master integrals forming the basis B_2:

$$\begin{array}{llll} \mathcal{J}_{T_2}(0,0,1,0,0,1,0), & \mathcal{J}_{T_2}(0,0,1,0,0,1,1), & \mathcal{J}_{T_2}(0,0,1,0,1,1,1), & \mathcal{J}_{T_2}(0,0,1,1,0,1,0), \\ \mathcal{J}_{T_2}(0,0,1,1,1,1,0), & \mathcal{J}_{T_2}(0,0,2,0,0,1,1), & \mathcal{J}_{T_2}(0,1,0,1,0,1,1), & \mathcal{J}_{T_2}(0,1,0,2,0,1,1), \\ \mathcal{J}_{T_2}(0,1,1,0,1,1,1), & \mathcal{J}_{T_2}(0,1,1,1,0,1,1), & \mathcal{J}_{T_2}(0,2,0,1,0,1,1), & \mathcal{J}_{T_2}(0,2,1,1,0,1,1), \\ \mathcal{J}_{T_2}(1,0,1,1,0,1,0), & \mathcal{J}_{T_2}(1,0,1,1,1,1,0), & & \end{array} \tag{D.3}$$

which we could express in linear combinations of the following 14 known masters:

$$\begin{array}{lllllll} \mathrm{Int}_b, & \mathrm{Int}_c, & \mathrm{Int}_d, & \mathrm{Int}_f, & \mathrm{Int}_g, & \mathrm{Int}_h, & \mathrm{Int}_i, \\ \mathrm{Int}_j, & \mathrm{Int}_k, & \mathrm{Int}_{\mathrm{mbmb}}, & \mathrm{Int}_{\mathrm{mbtri}}, & \mathrm{Int}_{\mathrm{tmb}}, & \mathrm{Int}_{\mathrm{tt}}, & \mathrm{Int}_{\mathrm{ttri}}. \end{array} \tag{D.4}$$

Basis for T_3 All the diagrams whose integrals fall in T_3 can be expressed as a linear combination of 11 master integrals forming the basis B_3:

$$\begin{array}{lll} \mathcal{J}_{\mathrm{T3}}(0,0,0,0,1,1,0) & \mathcal{J}_{\mathrm{T3}}(0,1,1,0,1,0,0) & \mathcal{J}_{\mathrm{T3}}(0,1,1,0,1,1,0) \\ \mathcal{J}_{\mathrm{T3}}(0,2,1,0,1,0,0) & \mathcal{J}_{\mathrm{T3}}(1,0,1,0,1,0,0) & \mathcal{J}_{\mathrm{T3}}(1,0,1,0,1,0,1) \\ \mathcal{J}_{\mathrm{T3}}(1,0,1,0,2,0,1) & \mathcal{J}_{\mathrm{T3}}(1,0,1,1,1,0,1) & \mathcal{J}_{\mathrm{T3}}(1,1,1,1,1,0,1) \\ \mathcal{J}_{\mathrm{T3}}(2,0,1,0,1,0,1) & \mathcal{J}_{\mathrm{T3}}(2,0,1,1,1,0,1) & 0 \end{array} \tag{D.5}$$

which we could express in linear combinations of the following 11 known masters:

$$\mathrm{Int}_a,\ \mathrm{Int}_b,\ \mathrm{Int}_c,\ \mathrm{Int}_f,\ \mathrm{Int}_g,\ \mathrm{Int}_h,\ \mathrm{Int}_i,\ \mathrm{Int}_j,\ \mathrm{Int}_k,\ \mathrm{Int}_{\mathrm{tmb}},\ \mathrm{Int}_{\mathrm{tt}}. \tag{D.6}$$

Basis rotation for T_1

$$
\begin{aligned}
\mathcal{J}_{T_1}(0,0,0,0,0,1,1) =&\quad \mathrm{Int}_{\mathrm{tt}}\\
\mathcal{J}_{T_1}(0,0,0,1,0,1,1) =&\quad \mathrm{Int}_{\mathrm{tmb}}\\
\mathcal{J}_{T_1}(0,0,0,1,1,1,1) =&\quad \mathrm{Int}_{\mathrm{ttri}}\\
\mathcal{J}_{T_1}(0,0,1,0,0,1,1) =&\quad \mathrm{Int}_{j}\\
\mathcal{J}_{T_1}(0,0,1,0,1,1,1) =&\quad \mathrm{Int}_{f}\\
\mathcal{J}_{T_1}(0,0,2,0,0,1,1) =&\quad \mathrm{Int}_{j}\frac{\left(\left(-7\mathfrak{s}^2+14\mathfrak{s}-16\right)d^2+\left(37\mathfrak{s}^2-74\mathfrak{s}+88\right)d-24\left(2\mathfrak{s}^2-4\mathfrak{s}+5\right)\right)}{(d-4)(\mathfrak{s}-4)^2\mathfrak{s}}\\
&+\mathrm{Int}_{k}\frac{24\left(d^2-5d+6\right)(\mathfrak{s}-1)}{(d-4)(\mathfrak{s}-4)^2\mathfrak{s}}+\mathrm{Int}_{\mathrm{tt}}\frac{(d-2)(-22\mathfrak{s}+d(7\mathfrak{s}-4)+16)}{(d-4)(\mathfrak{s}-4)^2\mathfrak{s}}\\
\mathcal{J}_{T_1}(0,1,0,1,0,1,1) =&\quad \mathrm{Int}_{g}\\
\mathcal{J}_{T_1}(0,1,0,2,0,1,1) =&\quad -\mathrm{Int}_{\mathrm{tt}}\frac{(-4\mathfrak{s}+d(\mathfrak{s}+2)-6)(d-2)^2}{4(d-3)(2d-7)(\mathfrak{s}-4)\mathfrak{s}}\\
&+\mathrm{Int}_{g}\frac{\left(\left(\mathfrak{s}^2-2\mathfrak{s}+4\right)d^2+\left(-7\mathfrak{s}^2+15\mathfrak{s}-24\right)d+4\left(3\mathfrak{s}^2-7\mathfrak{s}+9\right)\right)}{(2d-7)(\mathfrak{s}-4)\mathfrak{s}}\\
&-\mathrm{Int}_{h}\frac{(3d-8)(-4\mathfrak{s}+d(\mathfrak{s}+2)-6)}{(2d-7)(\mathfrak{s}-4)\mathfrak{s}}-\mathrm{Int}_{i}\frac{2(\mathfrak{s}-2)}{(2d-7)(\mathfrak{s}-4)}\\
&+\mathrm{Int}_{\mathrm{tmb}}\frac{(d-3)(3\mathfrak{s}d+2d-10\mathfrak{s}-8)}{2(2d-7)(\mathfrak{s}-4)\mathfrak{s}}\\
\mathcal{J}_{T_1}(0,2,0,1,0,1,1) =&\quad \mathrm{Int}_{g}\frac{(\mathfrak{s}+2)(d-3)^2}{(2d-7)\mathfrak{s}}+\mathrm{Int}_{\mathrm{tmb}}\frac{(d-4)(d-3)}{2(2d-7)\mathfrak{s}}+\mathrm{Int}_{h}\frac{\left(3d^2-17d+24\right)}{7\mathfrak{s}-2d\mathfrak{s}}\\
&+\mathrm{Int}_{i}\frac{(20-6d)}{2d^2-15d+28}+\mathrm{Int}_{\mathrm{tt}}\frac{\left(d^2-4d+4\right)}{28\mathfrak{s}-8d\mathfrak{s}}\\
\mathcal{J}_{T_1}(1,0,1,0,0,0,1) =&\quad \mathrm{Int}_{\mathrm{tb}}\\
\mathcal{J}_{T_1}(1,0,1,1,0,1,0) =&\quad \mathrm{Int}_{\mathrm{bmb}}\\
\mathcal{J}_{T_1}(1,0,1,1,0,1,1) =&\quad \mathrm{Int}_{e},
\end{aligned}
$$

Basis rotation for T_2

$$
\begin{aligned}
\mathcal{J}_{T_2}(0,0,1,0,0,1,0) \to\ & \mathrm{Int}_{\mathrm{tt}} \\
\mathcal{J}_{T_2}(0,0,1,0,0,1,1) \to\ & \mathrm{Int}_{j} \\
\mathcal{J}_{T_2}(0,0,1,0,1,1,1) \to\ & \mathrm{Int}_{f} \\
\mathcal{J}_{T_2}(0,0,1,1,0,1,0) \to\ & \mathrm{Int}_{\mathrm{tmb}} \\
\mathcal{J}_{T_2}(0,0,1,1,1,1,0) \to\ & \mathrm{Int}_{\mathrm{ttri}} \\
\mathcal{J}_{T_2}(0,0,2,0,0,1,1) \to\ & \tfrac{(2(\mathfrak{s}+3)-d(\mathfrak{s}+2))}{2(\mathfrak{s}-4)}\mathrm{Int}_{j} + \tfrac{3(d-2)}{\mathfrak{s}-4}\mathrm{Int}_{k} + \tfrac{(d-2)}{\mathfrak{s}-4}\mathrm{Int}_{\mathrm{tt}} \\
\mathcal{J}_{T_2}(0,1,0,1,0,1,1) \to\ & \mathrm{Int}_{g} \\
\mathcal{J}_{T_2}(0,1,0,2,0,1,1) \to\ & \tfrac{\left(d^2\left(\mathfrak{s}^2-2\mathfrak{s}+4\right)+d\left(-7\mathfrak{s}^2+15\mathfrak{s}-24\right)+4\left(3\mathfrak{s}^2-7\mathfrak{s}+9\right)\right)}{(2d-7)(\mathfrak{s}-4)\mathfrak{s}}\mathrm{Int}_{g} - \tfrac{(3d-8)(d(\mathfrak{s}+2)-4\mathfrak{s}-6)}{(2d-7)(\mathfrak{s}-4)\mathfrak{s}}\mathrm{Int}_{h} \\
& - \tfrac{2(\mathfrak{s}-2)}{(2d-7)(\mathfrak{s}-4)}\mathrm{Int}_{i} + \tfrac{(d-3)(3d\mathfrak{s}+2d-10\mathfrak{s}-8)}{2(2d-7)(\mathfrak{s}-4)\mathfrak{s}}\mathrm{Int}_{\mathrm{tmb}} - \tfrac{(d-2)^2(d(\mathfrak{s}+2)-4\mathfrak{s}-6)}{4(d-3)(2d-7)(\mathfrak{s}-4)\mathfrak{s}}\mathrm{Int}_{\mathrm{tt}} \\
\mathcal{J}_{T_2}(0,1,1,0,1,1,1) \to\ & \mathrm{Int}_{d} \\
\mathcal{J}_{T_2}(0,1,1,1,0,1,1) \to\ & \mathrm{Int}_{b} \\
\mathcal{J}_{T_2}(0,2,0,1,0,1,1) \to\ & \tfrac{\left(3d^2-17d+24\right)}{7\mathfrak{s}-2d\mathfrak{s}}\mathrm{Int}_{h} - \tfrac{\left(d^2-7d+12\right)}{14\mathfrak{s}-4d\mathfrak{s}}\mathrm{Int}_{\mathrm{tmb}} + \tfrac{\left(d^2-4d+4\right)}{28\mathfrak{s}-8d\mathfrak{s}}\mathrm{Int}_{\mathrm{tt}} \\
& + \tfrac{(d-3)^2(\mathfrak{s}+2)}{(2d-7)\mathfrak{s}}\mathrm{Int}_{g} + \tfrac{2}{2d-7}\mathrm{Int}_{i} \\
\mathcal{J}_{T_2}(0,2,1,1,0,1,1) \to\ & -\tfrac{1}{2}(d-4)^2\mathrm{Int}_{b} + \tfrac{2\left(d^2-7d+12\right)}{\mathfrak{s}}\mathrm{Int}_{c} + \tfrac{\left(d^2-7d+12\right)}{2\mathfrak{s}}\mathrm{Int}_{f} \\
& - \tfrac{\left(d^2-7d+12\right)((d-3)\mathfrak{s}+1)}{(2d-7)\mathfrak{s}}\mathrm{Int}_{g} + \tfrac{3\left(d^2-7d+10\right)}{(\mathfrak{s}-4)\mathfrak{s}}\mathrm{Int}_{k} + \tfrac{\left(-3d^3+29d^2-92d+96\right)}{7\mathfrak{s}-2d\mathfrak{s}}\mathrm{Int}_{h} \\
& + \tfrac{(d-2)\left(2d^3(\mathfrak{s}-4)+d^2(76-15\mathfrak{s})+d(35\mathfrak{s}-276)-26\mathfrak{s}+384\right)}{8(2d-7)(\mathfrak{s}-4)\mathfrak{s}}\mathrm{Int}_{\mathrm{tt}} \\
& + \tfrac{\left(d^4-10d^3+33d^2-37d+4\right)}{7\mathfrak{s}-2d\mathfrak{s}}\mathrm{Int}_{\mathrm{tmb}} + \tfrac{2(d-3)}{2d-7}\mathrm{Int}_{i} - \tfrac{(d-5)(d(\mathfrak{s}+2)-2(\mathfrak{s}+3))}{2(\mathfrak{s}-4)\mathfrak{s}}\mathrm{Int}_{j} \\
\mathcal{J}_{T_2}(1,0,1,1,0,1,0) \to\ & \mathrm{Int}_{\mathrm{mb2}} \\
\mathcal{J}_{T_2}(1,0,1,1,1,1,0) \to\ & \mathrm{Int}_{\mathrm{mbtri}}
\end{aligned}
$$

Basis rotation for T_3

$$
\begin{aligned}
\mathcal{J}_{T_3}(0,0,0,0,1,1,0) \to\ & \mathrm{Int}_{\mathrm{tt}} \\
\mathcal{J}_{T_3}(0,1,1,0,1,0,0) \to\ & \mathrm{Int}_{j} \\
\mathcal{J}_{T_3}(0,1,1,0,1,1,0) \to\ & \mathrm{Int}_{f} \\
\mathcal{J}_{T_3}(0,2,1,0,1,0,0) \to\ & \tfrac{\left(-7\mathfrak{s}^2+14\mathfrak{s}-16\right)d^2+\left(37\mathfrak{s}^2-74\mathfrak{s}+88\right)d-24\left(2\mathfrak{s}^2-4\mathfrak{s}+5\right)}{(d-4)(\mathfrak{s}-4)^2\mathfrak{s}}\mathrm{Int}_{j} \\
& +\tfrac{24\left(d^2-5d+6\right)(\mathfrak{s}-1)}{(d-4)(\mathfrak{s}-4)^2\mathfrak{s}}\mathrm{Int}_{k}+\tfrac{(d-2)(-22\mathfrak{s}+d(7\mathfrak{s}-4)+16)}{(d-4)(\mathfrak{s}-4)^2\mathfrak{s}}\mathrm{Int}_{\mathrm{tt}} \\
\mathcal{J}_{T_3}(1,0,1,0,1,0,0) \to\ & \mathrm{Int}_{\mathrm{tmb}} \\
\mathcal{J}_{T_3}(1,0,1,0,1,0,1) \to\ & \mathrm{Int}_{g} \\
\mathcal{J}_{T_3}(1,0,1,0,2,0,1) \to\ & \tfrac{3d^2-17d+24}{7\mathfrak{s}-2d\mathfrak{s}}\mathrm{Int}_{h}-\left(\tfrac{d^2-7d+12}{14\mathfrak{s}-4d\mathfrak{s}}\right)\mathrm{Int}_{\mathrm{tmb}}+\tfrac{d^2-4d+4}{28\mathfrak{s}-8d\mathfrak{s}}\mathrm{Int}_{\mathrm{tt}} \\
& +\tfrac{(d-3)^2(\mathfrak{s}+2)}{(2d-7)\mathfrak{s}}\mathrm{Int}_{g}+\tfrac{2}{2d-7}\mathrm{Int}_{i} \\
\mathcal{J}_{T_3}(1,0,1,1,1,0,1) \to\ & \mathrm{Int}_{b} \\
\mathcal{J}_{T_3}(1,1,1,1,1,0,1) \to\ & \mathrm{Int}_{a} \\
\mathcal{J}_{T_3}(2,0,1,0,1,0,1) \to\ & \tfrac{\left(\mathfrak{s}^2-2\mathfrak{s}+4\right)d^2+\left(-7\mathfrak{s}^2+15\mathfrak{s}-24\right)d+4\left(3\mathfrak{s}^2-7\mathfrak{s}+9\right)}{(2d-7)(\mathfrak{s}-4)\mathfrak{s}}\mathrm{Int}_{g}-\left(\tfrac{(3d-8)(-4\mathfrak{s}+d(\mathfrak{s}+2)-6)}{(2d-7)(\mathfrak{s}-4)\mathfrak{s}}\right)\mathrm{Int}_{h} \\
& -\left(\tfrac{2(\mathfrak{s}-2)}{(2d-7)(\mathfrak{s}-4)}\right)\mathrm{Int}_{i}+\tfrac{(d-3)(3\mathfrak{s}d+2d-10\mathfrak{s}-8)}{2(2d-7)(\mathfrak{s}-4)\mathfrak{s}}\mathrm{Int}_{\mathrm{tmb}} \\
& -\left(\tfrac{(d-2)^2(-4\mathfrak{s}+d(\mathfrak{s}+2)-6)}{4(d-3)(2d-7)(\mathfrak{s}-4)\mathfrak{s}}\right)\mathrm{Int}_{\mathrm{tt}} \\
\mathcal{J}_{T_3}(2,0,1,1,1,0,1) \to\ & -\mathrm{Int}_{c}\left(\tfrac{4(d-3)}{(\mathfrak{s}-4)\mathfrak{s}}\right)\left(\tfrac{(d-2)\left((7\mathfrak{s}+20)d^2-2(23\mathfrak{s}+64)d+76\mathfrak{s}+200\right)}{4(d-3)(2d-7)(\mathfrak{s}-4)^2\mathfrak{s}}\right)\mathrm{Int}_{\mathrm{tt}} \\
& +\tfrac{4d^3-25d^2+43d-16}{2(2d-7)(\mathfrak{s}-4)\mathfrak{s}}\mathrm{Int}_{\mathrm{tmb}}+\tfrac{d(\mathfrak{s}-4)-2(\mathfrak{s}-6)}{2(\mathfrak{s}-4)\mathfrak{s}}\mathrm{Int}_{f}+\tfrac{(d-3)(d(\mathfrak{s}-2)-3\mathfrak{s}+8)}{(2d-7)(\mathfrak{s}-4)\mathfrak{s}}\mathrm{Int}_{g} \\
& +\left(-\tfrac{(d-3)(3d-8)}{(2d-7)(\mathfrak{s}-4)\mathfrak{s}}\right)\mathrm{Int}_{h}+\tfrac{2}{(2d-7)(\mathfrak{s}-4)}\mathrm{Int}_{i}+\tfrac{d\left(\mathfrak{s}^2+\mathfrak{s}+16\right)-2\left(\mathfrak{s}^2+\mathfrak{s}+22\right)}{2(\mathfrak{s}-4)^2\mathfrak{s}}\mathrm{Int}_{j} \\
& +\left(-\tfrac{3(d-2)(\mathfrak{s}+2)}{(\mathfrak{s}-4)^2\mathfrak{s}}\right)\mathrm{Int}_{k}
\end{aligned}
$$

D.3 Renormalization scheme

Let us give the precise definition of our renormalization scheme.

- The gluon field is renormalized on-shell
- The strong coupling is renormalized in $\overline{\mathrm{MS}}$ with the top quark effects decoupled from the running by subtracting them at 0 momentum.
- The top mass corrections are subtracted in $\overline{\mathrm{MS}}$. The corrections generated by C_3 are decoupled from the running by multiplying them by $(\mu^2/m_t^2)^\epsilon$ (which yields a piece of the on-shell mass).
- The Wilson coefficients are renormalized in $\overline{\mathrm{MS}}$

This scheme choice ensures that the running of α_s and the top mass are exactly identical to the SM running with five light flavors. To keep expressions simple in the analytic calculation, we have not used the extra finite pieces from the on-shell top mass which would give the same result for the running.

D.4 Two-loop amplitudes

Here we present the results for the two-loop subtracted amplitudes $\mathcal{M}_1^F$ and $\mathcal{M}_3^F$ which we described in Section 5.1. Note that $\mathcal{M}_1^{\text{SM}}$ is proportional to $\mathcal{M}_1^F$.

$$
\begin{aligned}
\mathcal{M}_1^F = -\frac{ix}{90\sqrt{2}(x-1)^5 m_t}\Bigg(& \\
& G(0,x)\left(-180(x+1)(x-1)^2\log\left(\frac{\mu^2}{m_t^2}\right) - 10(x+1)\left(2\pi^2\left(x^2-1\right)\right.\right. \\
& \left.\left. -3\left(x^2(43\zeta(3)+4) - 8x - 11\zeta(3) + 4\right)\right) - 90\beta_0(x-1)^3\right) \\
& + G(0,0,x)\left(45(x-1)\left(4\left(x^2+6x+1\right) + \beta_0(x+1)^2\right)\log\left(\frac{\mu^2}{m_t^2}\right)\right. \\
& \left. + 5\left(75x^3 + 411x^2 + \pi^2\left(17x^3+17x^2-x-1\right) - 411x - 75\right)\right) \\
& - 30(x-1)\left(-47x^2 + 3\beta_0(x+1)^2 - 66x - 47\right)G(1,0,0,x) \\
& + 15\left(9\beta_0(x-1)(x+1)^2 - 2\left(33x^3+58x^2+59x+6\right)\right)G(0,0,0,x) \\
& - 480\left(x^3+x^2+x+1\right)G(0,-1,0,0,x) + 60\left(25x^3+25x^2+7x+7\right)G(0,0,-1,0,x) \\
& + 15\left(13x^3+13x^2-5x-5\right)G(0,0,0,0,x) + 240\left(x^3+x^2+x+1\right)G(0,0,1,0,x) \\
& - 840\left(x^3+x^2+x+1\right)G(0,1,0,0,x) - 90\beta_0(x+1)^2(x-1)G(0,0,1,x) \\
& - 30\left(3\beta_0+4\right)(x+1)^2(x-1)G(0,1,0,x) + 180(x-1)G(1,x)\left(\beta_0(x-1)^2 - 9(x+1)^2\zeta(3)\right) \\
& - 90\pi^2(x+1)^2(x-1)G(1,0,x) - 480(x+1)^2(x-1)G(0,-1,0,x) \\
& - 1080(x+1)^2(x-1)G(1,0,-1,0,x) + 540(x+1)^2(x-1)G(1,0,0,0,x) \\
& - 90\left(\beta_0+4\right)(x-1)^3\log\left(\frac{\mu^2}{m_t^2}\right) \\
& + 6\left(5(x-1)\left(x^2(31\zeta(3)-63) + 2x(17\zeta(3)+63) + 31\zeta(3) - 63\right) + \pi^4\left(3x^3+3x^2-x-1\right)\right)\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{M}_3^F = &-\frac{im_t}{540\sqrt{2}(x-1)^5(x+1)}\Big(\\
&5760(x+1)G(0,-1,0,x)(x-1)^5 - 180(x+1)\log^2\left(\frac{\mu^2}{m_t^2}\right)(6\beta_0+5)(x-1)^5 \\
&-5760(x+1)^2G(-1,0,x)(x-1)^4 - 1080(x+1)^2G(0,1,x)\beta_0(x-1)^4 \\
&-5760\left(x^3+x^2+x+1\right)G(-1,0,0,x)(x-1)^3 - 25920x(x+1)G(1,0,-1,0,x)(x-1)^3 \\
&+14940x(x+1)G(1,0,0,0,x)(x-1)^3 - 1800x(x+1)G(1,0,1,0,x)(x-1)^3 \\
&-2160x(x+1)G(1,1,0,0,x)(x-1)^3 - 2160x(x+1)G(0,0,1,x)\beta_0(x-1)^3 \\
&-180(x+1)\log\left(\frac{\mu^2}{m_t^2}\right)\left(3\beta_0(x-1)^2+62x^2-76x+62\right)(x-1)^3 \\
&-60(x+1)G(1,0,x)\left(31\pi^2x+27\left(x^2-1\right)+18\left(x^2-1\right)\beta_0\right)(x-1)^3 \\
&-720(x+1)G(1,0,0,x)\left(4x^2+(3\beta_0-71)x+4\right)(x-1)^3 \\
&+360(x+1)G(0,1,0,x)\left(8x^2-(6\beta_0+1)x+8\right)(x-1)^3 - 11520x\left(x^2+1\right)G(0,-1,0,0,x)(x-1)^2 \\
&+1440x\left(25x^2+7\right)G(0,0,-1,0,x)(x-1)^2 + 720x\left(7x^2-3\right)G(0,0,0,0,x)(x-1)^2 \\
&+180x\left(37x^2+27\right)G(0,0,1,0,x)(x-1)^2 - 360x\left(53x^2+59\right)G(0,1,0,0,x)(x-1)^2 \\
&+\Big(-210\pi^2\left(x^2-1\right)^2+180\left((-238+24\zeta(3))x^2+3(135+64\zeta(3))x+24\zeta(3)-238\right)\left(x^2-1\right) \\
&+\pi^4x\left(395x^2-107\right)\Big)(x-1)^2 + 360x(x+1)G(0,0,0,x)\left(9\beta_0(x-1)^3-152x^3+117x^2-210x-43\right) \\
&+G(0,0,x)\Bigg(360x\left(x^2-1\right)\log\left(\frac{\mu^2}{m_t^2}\right)\left(3\beta_0(x-1)^2+29x^2-10x+29\right) \\
&+30(x-1)\left(114x^5+3\left(368+21\pi^2\right)x^4-3\left(68+21\pi^2\right)x^3+\left(288+\pi^2\right)x^2\right. \\
&\left.-\left(-1350+\pi^2\right)x+36(x-1)^3(x+1)^2\beta_0-132\right)\Bigg) \\
&+G(1,x)\left(2160(x+1)\log\left(\frac{\mu^2}{m_t^2}\right)\beta_0(x-1)^5+360(x+1)\left(3(x-1)^2\beta_0-104x\zeta(3)\right)(x-1)^3\right) \\
&+G(0,x)\Big(90(x-1)^2\left(213x^4+\left(62-10\pi^2+336\zeta(3)\right)x^3-302x^2\right. \\
&\left.+2\left(31+5\pi^2-40\zeta(3)\right)x-6(x-1)^3(x+1)\beta_0+213\right) \\
&-180(x-1)^4(x+1)\log\left(\frac{\mu^2}{m_t^2}\right)(3(x-3)\beta_0-5(x+1))\Big)\Big)
\end{aligned}
$$

Bibliography

[1] S. L. Glashow, *Partial-symmetries of weak interactions*, *Nuclear Physics* **22** (1961), no. 4 579–588.

[2] S. Weinberg, *A Model of Leptons*, *Physical Review Letters* **19** (nov, 1967) 1264–1266.

[3] A. Salam, *Weak and Electromagnetic Interactions*, in *Conf. Proc. 8th Nobel Symposium Lerum* (N. Svartholm, ed.), vol. C680519, pp. 367–377, Almquist & Wiksell, 1968.

[4] A. H. Compton, *A Quantum Theory of the Scattering of X-rays by Light Elements*, *Physical Review* **21** (may, 1923) 483–502.

[5] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Physics Letters B* **716** (2012), no. 1 1–29, [arXiv:1207.7214].

[6] The CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Nuclear Physics B - Proceedings Supplements* **716** (jul, 2012) 30–61, [arXiv:1207.7235].

[7] The ATLAS Collaboration and The CMS Collaboration, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, *Journal of High Energy Physics* (2016), no. 8 45, [arXiv:1507.04548].

[8] Q. R. Ahmad, R. C. Allen, T. C. Andersen, et al., *Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory*, *Physical Review Letters* **89** (jun, 2002) 011301, [arXiv:nucl-ex/0204008].

[9] K. Eguchi, S. Enomoto, K. Furuno, et al., *First Results from KamLAND: Evidence for Reactor Antineutrino Disappearance*, *Physical Review Letters* **90** (jan, 2003) 021802, [arXiv:hep-ex/0212021].

[10] Y. Fukuda, T. Hayakawa, E. Ichihara, et al., *Evidence for Oscillation of Atmospheric Neutrinos*, *Physical Review Letters* **81** (aug, 1998) 1562–1567, [arXiv:hep-ex/9807003].

[11] F. Zwicky, *On the Masses of Nebulae and of Clusters of Nebulae*, *The Astrophysical Journal* **86** (oct, 1937) 217.

[12] A. Borriello and P. Salucci, *The dark matter distribution in disc galaxies*, *Monthly Notices of the Royal Astronomical Society* **323** (may, 2001) 285–292, [arXiv:astro-ph/0001082].

[13] N. A. Bahcall and X. Fan, *The Most Massive Distant Clusters: Determining Ω and σ_8*, *The Astrophysical Journal* **504** (1998), no. 1 1–6, [arXiv:astro-ph/9803277].

[14] J. A. Tyson, G. P. Kochanski, and I. P. Dell'Antonio, *Detailed Mass Map of CL 0024+1654 from Strong Lensing*, *The Astrophysical Journal* **498** (1998), no. 2 L107–L110, [arXiv:astro-ph/9801193].

[15] D. Clowe, M. Bradač, A. H. Gonzalez, et al., *A Direct Empirical Proof of the Existence of Dark Matter*, *The Astrophysical Journal* **648** (sep, 2006) L109–L113, [arXiv:astro-ph/0608407].

[16] P. A. R. Ade, N. Aghanim, M. Arnaud, et al., *Planck 2015 results*, *Astronomy and Astrophysics* **594** (2016) A13, [arXiv:1502.01589].

[17] M. E. Peskin and T. Takeuchi, *New constraint on a strongly interacting Higgs sector*, *Physical Review Letters* **65** (aug, 1990) 964–967.

[18] M. E. Peskin and T. Takeuchi, *Estimation of oblique electroweak corrections*, *Physical Review D* **46** (1992), no. 1 381–409.

[19] ATLAS Collaboration, *Measurement of inclusive jet and dijet cross sections in proton-proton collisions at 7 TeV centre-of-mass energy with the ATLAS detector*, *The European Physical Journal C* **71** (sep, 2011) 1512, [arXiv:1009.5908].

[20] The CMS Collaboration, *Measurement of the differential dijet production cross section in proton-proton collisions at $\sqrt{s}$ = 7 TeV*, *Physics Letters, Section B* **700** (2011), no. 3-4 187–206, [arXiv:1104.1693].

[21] ATLAS Collaboration, *Precision measurement and interpretation of inclusive W^+, W^- and Z/γ^* production cross sections with the ATLAS detector*, arXiv:1612.03016.

[22] CMS Collaboration, *Measurement of inclusive W and Z boson production cross sections in pp collisions at $\sqrt{s}$ = 8 TeV*, *Phys. Rev. Lett.* **112** (2014), no. 19 191802, [arXiv:1402.0923].

[23] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s}$=13 TeV with the ATLAS detector*, *Physics Letters B* **761** (jun, 2016) 136–157, [arXiv:1606.02699].

[24] The CMS Collaboration, *Measurement of the $t\bar{t}$ production cross section using events in the e μ final state in pp collisions at $\sqrt{s}$=13 TeV*, *European Physical Journal C* **77** (2017), no. 3 172, [arXiv:1611.04040].

[25] T. Hahn and M. Pérez-Victoria, *Automated one-loop calculations in four and D dimensions*, *Computer Physics Communications* **118** (may, 1999) 153–165, [arXiv:hep-ph/9807565].

[26] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *Journal of High Energy Physics* **2010** (jun, 2010) 43, [arXiv:1002.2581].

[27] G. Cullen, N. Greiner, G. Heinrich, et al., *Automated one-loop calculations with GoSam*, *The European Physical Journal C* **72** (mar, 2012) 1889, [arXiv:1111.2034].

[28] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *Automating the Powheg method in Sherpa*, *Journal of High Energy Physics* **2011** (apr, 2011) 24, [1008.5399v].

[29] G. Bevilacqua, M. Czakon, M. Garzelli, et al., *HELAC-NLO*, *Computer Physics Communications* **184** (mar, 2013) 986–997, [arXiv:1110.1499].

[30] S. Actis, A. Denner, L. Hofer, A. Scharf, and S. Uccirati, *Recursive generation of one-loop amplitudes in the Standard Model*, *Journal of High Energy Physics* **2013** (apr, 2013) 37, [arXiv:1211.6316].

[31] J. Alwall, R. Frederix, S. Frixione, et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *Journal of High Energy Physics* **2014** (jul, 2014) 79, [arXiv:1405.0301].

[32] S. Badger, B. Biedermann, P. Uwer, and V. Yundin, *Computation of multi-leg amplitudes with NJET*, *Journal of Physics: Conference Series* **523** (jun, 2014) 012057, [arXiv:1312.7140].

[33] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: The all-fermionic scattering channels*, *Journal of High Energy Physics* **2012** (2012), no. 12 [1207.0236v].

[34] O. Brein, A. Djouadi, and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, *Physics Letters B* **579** (2004), no. 1-2 149–156, [arXiv:hep-ph/0307206].

[35] P. Bolzoni, F. Maltoni, S.-O. Moch, and M. Zaro, *Vector boson fusion at next-to-next-to-leading order in QCD: Standard model Higgs boson and beyond*, *Physical Review D* **85** (feb, 2012) 035002, [arXiv:1109.3717].

[36] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini, *Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO*, *Physical Review Letters* **107** (mar, 2009) 152003, [arXiv:0903.2120].

[37] M. Grazzini, S. Kallweit, D. Rathlev, and M. Wiesemann, $W^{\pm}Z$ *production at hadron colliders in NNLO QCD*, *Physics Letters B* **761** (oct, 2016) 179–183, [arXiv:1604.08576].

[38] J. Currie, E. W. N. Glover, and J. Pires, *Next-to-Next-to Leading Order QCD Predictions for Single Jet Inclusive Production at the LHC*, *Physical Review Letters* **118** (feb, 2017) 072002, [arXiv:1611.01460].

[39] A. G.-D. Ridder, T. Gehrmann, E. W. N. Glover, A. Huss, and T. A. Morgan, *Precise QCD predictions for the production of a Z boson in association with a hadronic jet*, *Physical Review Letters* **117** (2015), no. 2 022001, [arXiv:1507.02850].

[40] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog, and B. Mistlberger, *Higgs boson gluon-fusion production in N3LO QCD*, arXiv:1503.06056.

[41] F. A. Dreyer and A. Karlberg, *Vector-boson fusion Higgs production at N^3LO in QCD*, arXiv:1606.00840.

[42] P. Bechtle, S. Heinemeyer, O. Stål, T. Stefaniak, and G. Weiglein, *Applying exclusion likelihoods from LHC searches to extended Higgs sectors*, *European Physical Journal C* **75** (2015), no. 9 1–24, [arXiv:1507.06706].

[43] T. Hahn, S. Heinemeyer, W. Hollik, H. Rzehak, and G. Weiglein, *FeynHiggs: A program for the calculation of MSSM Higgs-boson observables - Version 2.6.5*, *Computer Physics Communications* **180** (2009), no. 8 1426–1427.

[44] M. Wiesemann, R. Frederix, S. Frixione, et al., *Higgs production in association with bottom quarks*, *Journal of High Energy Physics* **2015** (feb, 2015) 132, [arXiv:1409.5301].

[45] The LHC Higgs Cross Section Working Group, S. Heinemeyer, C. Mariotti, et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, *2010 European School of High-Energy Physics* (jul, 2013) 230, [arXiv:1307.1347].

[46] M. Cacciari, F. A. Dreyer, A. Karlberg, G. P. Salam, and G. Zanderighi, *Fully Differential Vector-Boson-Fusion Higgs Production at Next-to-Next-to-Leading Order*, *Physical Review Letters* **115** (2015), no. 8 1–6, [arXiv:1506.02660].

[47] S. Dawson, C. Jackson, L. H. Orr, L. Reina, and D. Wackeroth, *Associated Higgs boson production with top quarks at the CERN Large Hadron Collider: NLO QCD corrections*, *Physical Review D* **68** (2003), no. 3 [arXiv:hep-ph/hep-ph/0106293].

[48] S. Dawson, L. H. Orr, L. Reina, and D. Wackeroth, *Associated Top Quark-Higgs Boson Production at the LHC*, *Physical Review D* **67** (nov, 2002) 071503, [arXiv:hep-ph/0211438].

[49] W. Beenakker, S. Dittmaier, M. Krämer, et al., *Higgs Radiation Off Top Quarks at the Tevatron and the LHC*, *Physical Review Letters* **87** (oct, 2001) 201805, [arXiv:hep-ph/hep-ph/0107081].

[50] W. Beenakker, S. Dittmaier, M. Krämer, et al., *NLO QCD corrections to production in hadron collisions*, *Nuclear Physics B* **653** (mar, 2003) 151–203, [arXiv:hep-ph/0211352].

[51] C. Anastasiou, S. Beerli, S. Bucherer, A. Daleo, and Z. Kunszt, *Two-loop amplitudes and master integrals for the production of a Higgs boson via a massive quark and a scalar-quark loop*, *Journal of High Energy Physics* **2007** (nov, 2006) 082–082, [arXiv:hep-ph/0611236].

[52] U. Aglietti, R. Bonciani, G. Degrassi, and A. Vicini, *Analytic results for virtual QCD corrections to Higgs production and decay*, *Journal of High Energy Physics* **2007** (jan, 2007) 021–021, [arXiv:hep-ph/0611266].

[53] T. Inami, T. Kubota, and Y. Okada, *Effective gauge theory and the effect of heavy quarks in Higgs boson decays*, *Zeitschrift für Physik C Particles and Fields* **18** (mar, 1983) 69–80.

[54] J. Ellis, M. Gaillard, D. Nanopoulos, and C. Sachrajda, *Is the mass of the Higgs boson about 10 GeV?*, *Physics Letters B* **83** (may, 1979) 339–344.

[55] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, *Remarks on Higgs-boson interactions with nucleons*, *Physics Letters B* **78** (1978), no. 4 443–446.

[56] R. Barbieri, *Looking beyond the standard model: the supersymmetric option*, *La Rivista Del Nuovo Cimento Series 3* **11** (apr, 1988) 1–45.

[57] H. P. Nilles, *Supersymmetry, supergravity and particle physics*, *Physics Reports* **110** (1984), no. 1-2 1–162.

[58] D. Marzocca, M. Serone, and J. Shu, *General composite Higgs models*, *Journal of High Energy Physics* **2012** (aug, 2012) 13, [arXiv:1205.0770].

[59] G. Ferretti, *Gauge theories of partial compositeness: scenarios for Run-II of the LHC*, *Journal of High Energy Physics* **2016** (jun, 2016) 107, [arXiv:1604.06467].

[60] S. Weinberg, *Phenomenological Lagrangians*, *Physica A* **96** (1979), no. 1-2 327–340.

[61] W. Buchmüller and D. Wyler, *Effective lagrangian analysis of new interactions and flavour conservation*, *Nuclear Physics B* **268** (may, 1986) 621–653.

[62] C. N. Leung, S. T. Love, and S. Rao, *Low-energy manifestations of a new interactions scale: Operator analysis*, *Zeitschrift für Physik C* **31** (1986), no. 3 433–437.

[63] A. Anselm and A. Johansen, *Can the electroweak θ-term be observable?*, *Nuclear Physics B* **412** (1994), no. 3 553–573, [arXiv:arXiv:hep-ph/9305271v1].

[64] J. E. Kim and G. Carosi, *Axions and the strong CP problem*, *Reviews of Modern Physics* **82** (mar, 2010) 557–601.

[65] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440–1443.

[66] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, *Physical Review Letters* **13** (1964), no. 16 508–509, [arXiv:10.1103/PhysRevLett.13.508].

[67] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, *Physical Review Letters* **13** (1964), no. 9 321–323.

[68] G. Hooft, *Renormalizable Lagrangians for massive Yang-Mills fields*, *Nuclear Physics B* **35** (dec, 1971) 167–188.

[69] K. A. Olive and Others, *Review of Particle Physics*, *Chin. Phys.* **C38** (2014) 90001.

[70] H. Lehmann, K. Symanzik, and W. Zimmermann, *Zur Formulierung quantisierter Feldtheorien*, *Il Nuovo Cimento* **1** (jan, 1955) 205–225.

[71] G. C. Wick, *The Evaluation of the Collision Matrix*, *Physical Review* **80** (oct, 1950) 268–272.

[72] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, *FeynRules 2.0 - A complete toolbox for tree-level phenomenology*, *Computer Physics Communications* **185** (2014), no. 8 2250–2300, [arXiv:1310.1921].

[73] G. t' Hooft and M. Veltman, *Regularization and Renormalizationof Gauge Fields*, *Nuclear Physics B* **44** (1972) 189–213, [1011.1669v].

[74] T. Kinoshita, *Mass Singularities of Feynman Amplitudes*, *Journal of Mathematical Physics* **3** (1962), no. 4 650–677.

[75] T. D. Lee and M. Nauenberg, *Degenerate Systems and Mass Singularities*, *Physical Review* **133** (1964), no. 6B B1549–B1562.

[76] J. C. Collins, D. E. Soper, and G. Sterman, *Factorization of Hard Processes in QCD*, arXiv:hep-ph/0409313.

[77] S. Alekhin, *Parton Distributions for LHC*, *Nuclear Physics B - Proceedings Supplements* **205-206** (2010), no. C 230–236, [arXiv:0901.0002].

[78] R. D. Ball, V. Bertone, S. Carrazza, et al., *Parton distributions for the LHC run II*, *Journal of High Energy Physics* **2015** (2015), no. 4 1–148, [arXiv:1410.8849].

[79] S. Dulat, T. J. Hou, J. Gao, et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, arXiv:1506.07443.

[80] ATLAS Collaboration, *Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*, *European Physical Journal C* **75** (2015), no. 10 476, [arXiv:1506.05669].

[81] The CMS Collaboration, *Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV*, *Physical Review D* **92** (2015) 012004, [arXiv:1411.3441].

[82] S. Bolognesi, Y. Gao, A. V. Gritsan, et al., *Spin and parity of a single-produced resonance at the LHC*, *Physical Review D - Particles, Fields, Gravitation and Cosmology* **86** (2012), no. 9 1–23, [arXiv:1208.4018].

[83] D. de Florian, C. Grojean, F. Maltoni, et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, arXiv:1610.07922.

[84] The CMS Collaboration, *Search for two Higgs bosons in final states containing two photons and two bottom quarks*, *Physical Review D* **94** (2016), no. 5 052012, [arXiv:1603.06896].

[85] The ATLAS Collaboration, *Measurement of fiducial differential cross sections of gluon-fusion production of Higgs bosons decaying to $WW^* \to e\nu\mu\nu$ with the ATLAS detector at $\sqrt{s} = 8$ TeV*, *Journal of High Energy Physics* (2016), no. 9 104, [1604.02997].

[86] A. V. Manohar, *Effective Field Theories*, arXiv:hep-ph/9508245.

[87] K. G. Chetyrkin, B. A. Kniehl, and M. Steinhauser, *Decoupling Relations to $O(\alpha_s^3)$ and their Connection to Low-Energy Theorems*, arXiv:hep-ph/9708255.

[88] Y. Schroder and M. Steinhauser, *Four-Loop Decoupling Relations for the Strong Coupling*, arXiv:hep-ph/0512058.

[89] C. Anastasiou and K. Melnikov, *Higgs boson production at hadron colliders in NNLO QCD*, *Nuclear Physics B* **646** (2002), no. 1-2 220–256, [arXiv:hep-ph/0207004].

[90] ATLAS Collaboration, *Search for New Phenomena in Dijet Mass and Angular Distributions from pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*, *Phys. Lett. B* **754** (2016), no. December 302, [arXiv:1512.01530].

[91] CMS Collaboration, *Search for narrow resonances decaying to dijets in proton-proton collisions at sqrt(s) = 13 TeV*, arXiv:1512.01224.

[92] The ATLAS collaboration, M. Aaboud, G. Aad, et al., *Search for resonances in diphoton events at √s=13 TeV with the ATLAS detector*, *Journal of High Energy Physics* **2016** (2016), no. 9 1–49, [arXiv:1606.03833].

[93] CMS Collaboration, *Search for heavy narrow dilepton resonances in pp collisions at sqrt(s) = 7 TeV and sqrt(s) = 8 TeV*, *Physics Letters B* **720** (dec, 2012) 63–82, [arXiv:1212.6175].

[94] CMS Collaboration, *Search for dark matter, extra dimensions, and unparticles in monojet events in proton-proton collisions at $\sqrt{s}$ = 8 TeV*, *European Physical Journal C* **75** (aug, 2014) 235, [arXiv:1408.3583].

[95] ATLAS Collaboration, *Search for new phenomena in events containing a same-flavour opposite-sign dilepton pair, jets, and large missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, *The European Physical Journal C* **77** (nov, 2016) 1–63, [arXiv:1611.05791].

[96] CMS Collaboration, *Search for physics beyond the standard model in events with two leptons, jets, and missing transverse momentum in pp collisions at $\sqrt{s}$ = 8 TeV*, arXiv:1502.06031.

[97] H. Baer, C. Chen, and A. Bartl, *Low energy supersymmetry phenomenology*, tech. rep., Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA, 1995.

[98] H.-C. Cheng, K. T. Matchev, and M. Schmaltz, *Bosonic supersymmetry? Getting fooled at the CERN LHC*, *Physical Review D* **66** (2002), no. 5 056006, [arXiv:hep-ph/0205314].

[99] A. Arbey, G. Cacciapaglia, A. Deandrea, and B. Kubik, *Dark Matter in a twisted bottle*, *Journal of High Energy Physics* **2013** (2013), no. 1 43, [arXiv:1210.0384].

[100] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, *The Hierarchy Problem and New Dimensions at a Millimeter*, *Physics Letters B* **429** (1998), no. 4 263–272, [arXiv:hep-ph/9803315].

[101] L. Randall and R. Sundrum, *A Large Mass Hierarchy from a Small Extra Dimension*, *Phys. Rev. Lett.* **83** (1999) 3370, [arXiv:hep-ph/9905221].

[102] G. Cacciapaglia, A. Deandrea, and N. Deutschmann, *Dark matter and localised fermions from spherical orbifolds?*, *Journal of High Energy Physics* **2016** (2016), no. 4 [arXiv:1601.00081].

[103] J. Alwall, P. C. Schuster, and N. Toro, *Simplified models for a first characterization of new physics at the LHC*, *Physical Review D* **79** (2009), no. 7 075020, [arXiv:0810.3921].

[104] The GAMBIT Collaboration, P. Athron, C. Balázs, et al., *A global fit of the MSSM with GAMBIT*, arXiv:1705.07917.

[105] T. Appelquist and J. Carazzone, *Infrared singularities and massive fields*, *Physical Review D* **11** (1975), no. 10 2856–2861.

[106] 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.

[107] A. Falkowski and K. Mimouni, *Model independent constraints on four-lepton operators*, *Journal of High Energy Physics* **2016** (2016), no. 2 1–28, [arXiv:1511.07434].

[108] L. Berthier and M. Trott, *Towards consistent Electroweak Precision Data constraints in the SMEFT*, *Journal of High Energy Physics* **2015** (2015), no. 5 [arXiv:1502.02570].

[109] L. Berthier, M. Bjørn, and M. Trott, *Incorporating doubly resonant W ± data in a global fit of SMEFT parameters to lift flat directions*, *Journal of High Energy Physics* **2016** (2016), no. 9 [arXiv:1606.06693].

[110] H. Bélusca-Maïto, A. Falkowski, D. Fontes, J. C. Romão, and J. P. Silva, *Higgs EFT for 2HDM and beyond*, *The European Physical Journal C* **77** (2017), no. 3 176, [arXiv:1611.01112].

[111] B. Grzadkowski, M. Iskrzyński, M. Misiaka, and J. Rosieka, *Dimension-six terms in the Standard Model Lagrangian*, *Journal of High Energy Physics* **2010** (2010), no. 10 0–16, [arXiv:1008.4884].

[112] 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*, *Journal of High Energy Physics* **2014** (2014), no. 4 159, [arXiv:1312.2014].

[113] G. F. Giudice, C. Grojean, A. Pomarol, and R. Rattazzi, *The Strongly-Interacting Light Higgs*, *Journal of High Energy Physics* (2007), no. 06 045, [arXiv:hep-ph/0703164].

[114] E. E. Jenkins, A. V. Manohar, and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators I: Formalism and lambda Dependence*, arXiv:1308.2627.

[115] E. E. Jenkins, A. V. Manohar, and M. Trott, *Renormalization group evolution of the Standard Model dimension six operators II: Yukawa dependence*, *Journal of High Energy Physics* **2014** (2014), no. 1 [arXiv:1310.4838].

[116] C. Zhang, *Effective field theory approach to top-quark decay at next-to-leading order in QCD*, *Physical Review D - Particles, Fields, Gravitation and Cosmology* **90** (2014), no. 1 1–34, [arXiv:1404.1264].

[117] C. Zhang, *Single top production at next-to-leading order in the Standard Model effective field theory*, *Phys. Rev. Lett.* **116** (2016) 162002, [arXiv:1601.06163].

[118] F. Maltoni, E. Vryonidou, and C. Zhang, *Higgs production in association with a top-antitop pair in the Standard Model Effective Field Theory at NLO in QCD*, *Journal of High Energy Physics* (jul, 2016) 123, [arXiv:1607.05330].

[119] O. B. 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*, *Journal of High Energy Physics* **2016** (2016), no. 5 [arXiv:1601.08193].

[120] C. Zhang and F. Maltoni, *Top-quark decay into Higgs boson and a light quark at next-to-leading order in QCD*, *Physical Review D* **88** (2013) 054005, [arXiv:1305.7386].

[121] C. Hartmann and M. Trott, *On one-loop corrections in the standard model effective field theory; the $\Gamma(h \to \gamma\gamma)$ case*, *Journal of High Energy Physics* **2015** (2015), no. 7 [arXiv:1505.02646].

[122] C. Hartmann and M. Trott, *Higgs decay to two photons at one-loop in the SMEFT*, *Phys. Rev. Lett.* (2015), no. 115 191801, [arXiv:1507.03568].

[123] M. Grazzini, A. Ilnicka, M. Spira, and M. Wiesemann, *BSM effects on the Higgs transverse-momentum spectrum in an EFT approach*, *Journal of High Energy Physics* (2017), no. 10 115, [arXiv:1612.00283].

[124] K. Mimasu, V. Sanz, and C. Williams, *Higher order QCD predictions for associated Higgs production with anomalous couplings to gauge bosons*, *Journal of High Energy Physics* **2016** (2016), no. 8 [arXiv:1512.02572].

[125] P. Nogueira, *Automatic Feynman Graph Generation*, *Journal of Computational Physics* **105** (1993), no. 2 279–289.

[126] T. Hahn, *Generating Feynman diagrams and amplitudes with FeynArts 3*, *Computer Physics Communications* **140** (2001), no. 3 418–431, [arXiv:hep-ph/0012260].

[127] J. Kuipers, T. Ueda, J. a. M. Vermaseren, and J. Vollinga, *FORM version 4.0*, *Computer Physics Communications* **184** (2013), no. 5 1453–1467, [arXiv:1203.6543].

[128] A. V. Smirnov and A. V. Petukhov, *The Number of Master Integrals is Finite*, *Letters in Mathematical Physics* **97** (2011), no. 1 37–44, [arXiv:1004.4199].

[129] F. V. Tkachov, *A theorem on analytical calculability of 4-loop renormalization group functions*, *Physics Letters B* **100** (1981), no. 1 65–68.

[130] K. G. Chetyrkin and F. V. Tkachov, *Integration by parts: The algorithm to calculate β-functions in 4 loops*, *Nuclear Physics B* **192** (1981), no. 1 159–204.

[131] S. Laporta and E. Remiddi, *The analytical value of the electron light-light graphs contribution to the muon (g-2) in QED*, *Physics Letters B* **301** (1993), no. 4 440–446, [arXiv:hep-ph/9602417].

[132] S. Laporta, *Calculation of master integrals by difference equations*, *Physics Letters B* **504** (2001), no. 1-2 188–194, [arXiv:hep-ph/0102033].

[133] V. a. Smirnov, *Analytic Tools for Feynman Integrals*, vol. 250. Springer, Heidelberg, 2012.

[134] A. V. Smirnov, *FIRE5: a C++ implementation of Feynman Integral REduction*, *Computer Physics Communications* **189** (2015), no. April 182–191, [arXiv:1408.2372].

[135] A. V. Smirnov, *Algorithm FIRE—Feynman Integral REduction*, *Journal of High Energy Physics* **2008** (2008), no. 10 107–107, [arXiv:0807.3243v].

[136] C. Anastasiou and A. Lazopoulos, *Automatic Integral Reduction for Higher Order Perturbative Calculations*, *Journal of High Energy Physics* **2004** (jul, 2004) 046–046, [arXiv:hep-ph/0404258].

[137] R. N. Lee, *Presenting LiteRed: a tool for the Loop InTEgrals REDuction*, arXiv:1212.2685.

[138] P. Maierhoefer, J. Usovitsch, and P. Uwer, *Kira - A Feynman Integral Reduction Program*, arXiv:1705.05610.

[139] T. Gehrmann, E. Weihs, L. Tancredi, and A. von Manteuffel, *The Two-Loop Master Integrals for $q\bar{q} \to VV$*, *Journal of High Energy Physics* (2014), no. 06 032, [arXiv:1404.4853].

[140] A. Kotikov, *Differential equations method. New technique for massive Feynman diagram calculation*, *Physics Letters B* **254** (1991), no. 1-2 158–164, [arXiv:hep-th/9711188].

[141] E. Remiddi, *Differential Equations for Feynman Graph Amplitudes*, *Nuovo Cim.* **A110** (1997) 19, [arXiv:hep-th/9711188].

[142] J. M. Henn, *Multiloop integrals in dimensional regularization made simple*, *Physical Review Letters* **110** (2013) 251601, [arXiv:1304.1806].

[143] A. V. Smirnov, "UF — Mathematica Package." http://science.sander.su/Tools-UF.htm.

[144] N. Nakanishi, *Graph theory and Feynman integrals*. Mathematics and its applications. Gordon and Breach, 1971.

[145] H. Cheng and T. T. Wu, *Expanding Protons: Scattering At High Energies*. MIT Press, Cambridge, USA, 1987.

[146] M. Beneke and V. a. Smirnov, *Asymptotic expansion of Feynman integrals near threshold*, *Nuclear Physics B* **522** (nov, 1997) 321–344, [arXiv:hep-ph/9711391].

[147] V. A. Smirnov, *Problems of the Strategy of Regions*, *Phys. Lett. B* **465** (1999) 226–234, [arXiv:hep-ph/9907471].

[148] A. Pak and A. Smirnov, *Geometric approach to asymptotic expansion of feynman integrals*, *European Physical Journal C* **71** (2011), no. 4 1–6, [arXiv:1011.4863].

[149] a. Pak and A. Smirnov, *Geometric approach to asymptotic expansion of feynman integrals*, *European Physical Journal C* **71** (2011), no. 4 1–6, [arXiv:1011.4863].

[150] T. Binoth and G. Heinrich, *An automatized algorithm to compute infrared divergent multi-loop integrals*, *Nuclear Physics B* **585** (2000), no. 3 741–759.

[151] T. Binoth and G. Heinrich, *Numerical evaluation of phase space integrals by sector decomposition*, *Nuclear Physics B* **693** (2004), no. 1-3 134–148, [arXiv:hep-ph/0402265].

[152] G. Heinrich, *Sector Decomposition*, *International Journal of Modern Physics A* **23** (2008), no. 10 1457–1486, [arXiv:0803.4177].

[153] T. Ueda and J. Fujimoto, *New implementation of the sector decomposition on FORM*, in *PoS ACAT08*, p. 120, 2008. [arXiv:0902.2656].

[154] C. Bogner and S. Weinzierl, *Resolution of singularities for multi-loop integrals*, *Computer Physics Communications* **178** (2008), no. 8 596–610, [arXiv:0709.4092].

[155] J. Gluza, K. Kajda, T. Riemann, and V. Yundin, *Numerical Evaluation of Tensor Feynman Integrals in Euclidean Kinematics*, *European Physical Journal C* **71** (2011) 1516, [arXiv:1010.1667].

[156] S. Borowka, G. Heinrich, S. Jones, et al., *SecDec-3.0: Numerical evaluation of multi-scale integrals beyond one loop*, *Computer Physics Communications* **196** (2015) 470–491, [arXiv:1502.06595].

[157] A. V. Smirnov, *FIESTA 4: optimized Feynman integral calculations with GPU support*, *Computer Physics Communications* **204** (2016) 189–199, [arXiv:1511.03614].

[158] S. Borowka, N. Greiner, G. Heinrich, et al., *Full top quark mass dependence in Higgs boson pair production at NLO*, *Journal of High Energy Physics* **2016** (oct, 2016) 107, [arXiv:1608.04798].

[159] C. Duhr, *Mathematical aspects of scattering amplitudes*, in *Journeys Through The Precision Frontier: Amplitudes For Colliders (TASI 2014)* (L. Dixon, F. Petriello, and T. Degrand, eds.), ch. 10, pp. 419–476. World Scientific, Singapore, 2015. [arXiv:1411.7538].

[160] E. Remiddi and J. A. M. Vermaseren, *Harmonic Polylogarithms*, *International Journal of Modern Physics A* **15** (2000), no. 5 725, [arXiv:hep-ph/9905237].

[161] S. Bloch and P. Vanhove, *The elliptic dilogarithm for the sunset graph*, *Journal of Number Theory* **148** (2015) 328–364, [arXiv:1309.5865].

[162] L. Adams, C. Bogner, A. Schweitzer, and S. Weinzierl, *The kite integral to all orders in terms of elliptic polylogarithms*, *Journal of Mathematical Physics* **57** (2016), no. 12 [arXiv:1607.01571].

[163] R. Bonciani, V. D. Duca, H. Frellesvig, et al., *Two-loop planar master integrals for Higgs $\to$ 3 partons with full heavy-quark mass dependence*, *Journal of High Energy Physics* (2016), no. 12 096, [1609.06685].

[164] A. B. Goncharov, *Multiple polylogarithms, cyclotomy and modular complexes*, *Math Res. Letters* **5** (1998), no. 3 497–516, [arXiv:1105.2076].

[165] A. B. Goncharov, *Multiple polylogarithms and mixed Tate motives*, arXiv:math/0103059.

[166] A. B. Goncharov, *Galois symmetries of fundamental groupoids and noncommutative geometry*, *Duke Mathematical Journal* **128** (2005), no. 2 209–284, [arXiv:math/0208144].

[167] D. E. Radford, *A Natural Ring Basis fo the Shuffle Algebra and an Application to Group Schemes*, *Journal of Algebra* **58** (1979) 432–454.

[168] C. Duhr, *Hopf algebras, coproducts and symbols: An application to Higgs boson amplitudes*, *Journal of High Energy Physics* **2012** (2012), no. 8 [arXiv:1203.0454].

[169] F. Brown, *On the decomposition of motivic multiple zeta values*, arXiv:1102.1310.

[170] E. Panzer, *Feynman integrals and hyperlogarithms*, arXiv:1506.07243.

[171] C. Anastasiou, C. Duhr, F. Dulat, and B. Mistlberger, *Soft triple-real radiation for Higgs production at N3LO*, *Journal of High Energy Physics* **2013** (2013), no. 7 [arXiv:1302.4379].

[172] H. R. P. Ferguson, D. H. Bailey, and S. Arno, *Analysis of PSLQ, an integer relation finding algorithm*, *Mathematics of Computation* **68** (1999), no. 225 351–370.

[173] The CMS Collaboration, *Search for the standard model Higgs boson produced through vector boson fusion and decaying to bb*, *Physical Review D* **92** (2015) 032008, [arXiv:1506.01010].

[174] The ATLAS Collaboration, *Search for the Standard Model Higgs boson produced by vector-boson fusion and decaying to bottom quarks in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector*, *Journal of High Energy Physics* (2016), no. 11 112, [1606.02181].

[175] The CMS Collaboration, *Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks*, *Physical Review D* **89** (2014) 012003, [arXiv:1310.3687].

[176] ATLAS Collaboration, *Search for the standard model Higgs boson produced in association with a vector boson and decaying into a tau pair in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *Physical Review D* **93** (2016), no. 9 092005, [arXiv:1511.08352].

[177] H. Mantler and M. Wiesemann, *Top- and bottom-mass effects in hadronic Higgs production at small transverse momenta through lo+nll*, *European Physical Journal C* **73** (2013), no. 6 1–11, [arXiv:1210.8263].

[178] M. Grazzini and H. Sargsyan, *Heavy-quark mass effects in Higgs boson production at the LHC*, *Journal of High Energy Physics* **2013** (2013), no. 9 [arXiv:1306.4581].

[179] A. Banfi, P. F. Monni, and G. Zanderighi, *Quark masses in Higgs production with a jet veto*, *Journal of High Energy Physics* **2014** (2014), no. 1 97, [arXiv:1308.4634].

[180] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini, *Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM*, *Journal of High Energy Physics* **2012** (2012), no. 2 88, [arXiv:1111.2854].

[181] D. Rainwater, M. Spira, and D. Zeppenfeld, *Higgs Boson Production at Hadron Colliders: Signal and Background Processes*, in *Physics at TeV colliders. Proceedings, Euro Summer School, Les Houches.*, p. 15, mar, 2001. [arXiv:hep-ph/0203187].

[182] J. Campbell, S. Dawson, S. Dittmaier, et al., *Higgs Boson Production in Association with Bottom Quarks*, in *Physics at TeV colliders. Proceedings, Workshop, Les Houches.*, pp. 2–8, 2003. [arXiv:hep-ph/0405302].

[183] S. Dittmaier, M. Krämer, and M. Spira, *Higgs radiation off bottom quarks at the Fermilab Tevatron and the CERN LHC*, *Physical Review D* **70** (2004), no. 7 1–10, [arXiv:hep-ph/0309204].

[184] C. Duhr and F. Dulat, "Polylogtools." private code, 2014.

[185] D. B. Franzosi and C. Zhang, *Probing the top-quark chromomagnetic dipole moment at next-to-leading order in QCD*, *Physical Review D* **91** (2015), no. 11 1–11, [arXiv:1503.08841].

[186] R. Gauld, B. D. Pecjak, and D. J. Scott, *One-loop corrections to $h \to b\bar{b}$ and $h \to \tau^+\tau^-$ decays in the Standard Model dimension-6 EFT: four-fermion operators and the large-m_t limit*, *Journal of High Energy Physics* **2016** (2016), no. 5 [arXiv:1512.02508].

[187] C. Degrande, J.-M. Gérard, C. Grojean, F. Maltoni, and G. Servant, *Probing top-Higgs non-standard interactions at the LHC*, *Journal of High Energy Physics* **2012** (jul, 2012) 36, [arXiv:1205.1065].

[188] S. Catani and M. H. Seymour, *A General Algorithm for Calculating Jet Cross Sections in NLO QCD*, *Nuclear Physics B* **485** (1996), no. 1-2 291–419, [arXiv:hep-ph/9605323].

[189] S. Dawson, *Radiative corrections to Higgs boson production*, *Nuclear Physics B* **359** (1991), no. 2-3 283–300.

[190] R. J. Eden, P. V. Landshoff, D. I. Olive, and J. C. Polkinghorne, *The Analytic S-Matrix*. Cambridge University Press, Cambridge, 1966.

[191] H. Mantler and M. Wiesemann, *Hadronic Higgs production through NLO+ PS in the SM, the 2HDM and the MSSM*, *The European Physical Journal C* **75** (2015), no. 6 257, [arXiv:1504.06625].

[192] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, *European Physical Journal C* **75** (2015), no. 5 1–83, [arXiv:1412.3989].

[193] L. Berthier and M. Trott, *Consistent constraints on the Standard Model Effective Field Theory*, *Journal of High Energy Physics* **2016** (2016), no. 2 1–48, [arXiv:1508.05060].

[194] R. Contino, A. Falkowski, F. Goertz, C. Grojean, and F. Riva, *On the validity of the effective field theory approach to SM precision tests*, *Journal of High Energy Physics* **2016** (2016), no. 7 [1604.06444].

[195] G. Passarino and M. Trott, *The Standard Model Effective Field Theory and Next to Leading Order*, arXiv:1610.08356.

[196] N. Deutschmann, *Qgraf-XML-drawer 1.0*, https://doi.org/10.5281/zenodo.164393, 2016.

[197] J. Ellis, *TikZ-Feynman: Feynman diagrams with TikZ*, *Computer Physics Communications* **210** (2017) 103–123, [arXiv:1601.05437].