NLO QCD corrections for the $pp \to t\bar{t}H$ process in the dilepton decay channel within the SMEFT framework

Von der Fakultät für Mathematik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Minos Nikolaos Reinartz, M.Sc.
aus
Troisdorf

Berichter: apl. Prof. Dr. rer. nat. Malgorzata Worek
Univ.-Prof. Dr. rer. nat. Robert Harlander

Tag der mündlichen Prüfung: 18.09.2026

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek verfügbar.

Abstract

This thesis presents next-to-leading order (NLO) QCD corrections to the $pp \to t\bar{t}H + X$ process with dileptonic top-quark decays in the *Standard Model Effective Field Theory* (SMEFT) framework. Unstable particles are treated using the *narrow-width approximation* (NWA). The dimension-6 operators $\mathcal{O}_{t\phi}$, $\mathcal{O}_{\phi G}$, $\mathcal{O}_{tG}$ and $\mathcal{O}_{tW}$ are taken into account and are consistently included in both the production and decay matrix elements of the process. All results are provided at the LHC Run III center-of-mass energy of $\sqrt{s} = 13.6\,\text{TeV}$.

To facilitate this calculation, we present HELAC-SMEFT, a newly developed software package based on the existing HELAC-NLO program. The overhauled code can interface with existing literature results provided in the *Universal Feynman Output* format, treat unstable resonances consistently using either an off-shell approach or using the NWA and extends efficient reweighting methods to the use in SMEFT calculations. The new software is designed to study SMEFT at NLO QCD accuracy, but can also be used to compute Standard Model (SM) predictions. The platform provides a solid foundation for future extensions for both SM and SMEFT computations alike. We discuss the inner workings of the original HELAC-NLO code and present a detailed record of the newly implemented modifications, generalizations and optimizations in HELAC-SMEFT.

We provide a detailed analysis of linear and quadratic SMEFT contributions on both the integrated and differential level for the $pp \to t\bar{t}H \to (W^+b)(W^-\bar{b})H \to (e^+\nu_e b)(\mu^-\bar{\nu}_\mu \bar{b})H$ process. The findings are contextualized with a corresponding computation using stable top quarks. We study the size of NLO QCD corrections and theoretical uncertainties stemming from the truncation of the perturbative series as well as the *renormalization group* evolution of the Wilson coefficients. Similarities and differences of kinematical features between SMEFT corrections and the SM predictions are analysed for various differential observables. To evaluate the impact of the more realistic top-quark modelling approach, we reconstruct the top-quark kinematics from the respective decay products and investigate shape differences compared to the computation with stable top quarks. We show that phase-space cuts on the final state can have large effects on the shape of many LHC observables and that the size of these deviations can vary between the SM and various SMEFT contributions. Consequently, these effects have to be evaluated on a case by case basis.

We conclude that both higher order QCD corrections and realistic top-quark decay modelling are important for SMEFT analyses and should be incorporated when available.

Zusammenfassung

Diese Arbeit präsentiert QCD-Korrekturen der nächstführenden Ordnung (NLO) zum $pp \to t\bar{t}H + X$-Prozess mit dileptonischen Top-Quark-Zerfällen in der effektiven Feldtheorie des Standardmodells (SMEFT). Instabile Teilchen werden mit Hilfe der *Narrow-Width Approximation* (NWA) beschrieben. Die sechs-dimensionalen Operatoren $\mathcal{O}_{t\phi}$, $\mathcal{O}_{\phi G}$, $\mathcal{O}_{tG}$ und $\mathcal{O}_{tW}$ werden in den Matrixelementen des Produktions- und Zerfallsprozesses mit einbezogen. Alle Ergebnisse werden für die LHC-Run-III Schwerpunktsenergie von $\sqrt{s} = 13.6\,\text{TeV}$ berechnet.

Zur Durchführen dieser Berechnung präsentieren wir HELAC-SMEFT, ein neu entwickeltes Softwarepaket, welches auf dem HELAC-NLO-Programm basiert. Das überarbeitete Program besitzt eine Schnittstelle zu existierender Literatur im *Universal Feynman Output*-Format. Instabile Resonanzen können entweder mit einer *Off-shell*-Beschreibung oder der NWA konsistent behandelt werden. Effiziente Umgewichtungsmethoden wurden außerdem für Anwendungen in der SMEFT erweitert. Die neue Software ist auf SMEFT-Studien in nächstführender Ordnung in QCD spezialisiert, aber Standardmodell (SM) Vorhersagen können auch berechnet werden. Die Plattform bildet eine solide Grundlage für zukünftige Erweiterungen für weitere SM- und SMEFT- Berechnungen. Wir erläutern die Funktionsweise des ursprünglichen HELAC-NLO Programms und präsentieren eine detaillierte Dokumentations der im neuen Programm implementierten Modifikationen, Verallgemeinerungen und Optimierungen.

Wir präsentieren eine detaillierte Analyse von linearen und quadratischen SMEFT Beiträgen auf der integrierten und differenziellen Ebene für den Prozess $pp \to t\bar{t}H \to (W^+b)(W^-\bar{b})H \to (e^+\nu_e b)(\mu^-\bar{\nu}_\mu\bar{b})H$. Die Ergebnisse werden mit einer äquivalenten Berechnung mit stabilen Top-Quarks kontextualisiert. Wir studieren die Größe der NLO-QCD-Korrekturen und theoretischen Unsicherheiten, die einerseits von der Trunkierung der perturbativen Serie sowie von dem Einfluss von Renormalisierungsgruppeneffekten stammen. Ähnlichkeiten und Unterschiede kinematischer Merkmale zwischen den SMEFT-Korrekturen und dem SM-Vorhersagen werden für eine Vielzahl von Observablen analysiert. Um den Einfluss der realistischen Top-Quark-Modellierung zu quantifizieren, rekonstruieren wir die Top-Quark-Impulse aus den jeweiligen Zerfallsprodukten und evaluieren die Formunterschiede der Verteilungen im Vergleich zur Rechnung mit stabilen Top-Quarks. Wir zeigen, dass Restriktionen des Phasenraums des Endzustands große Effekte auf die Form vieler LHC-Observablen haben und dass sich die Größe dieser Abweichungen sich zwischen dem SM und den verschiedenen SMEFT-Korrekturen unterscheiden kann. Diese Effekte müssen daher von Fall zu Fall bewertet werden.

Wir schlussfolgern, dass sowohl nächstführende QCD-Korrekturen als auch eine realistische Beschreibung von Top-Quarks eine entscheidene Rolle für SMEFT-Analysen darstellen.

Eidesstattliche Erklärung

Minos Nikolaos Reinartz
erklärt hiermit, dass diese Dissertation und die darin dargelegten Inhalte die eigenen sind und selbstständig, als Ergebnis der eigenen originären Forschung, generiert wurden. Hiermit erkläre ich an Eides statt

1. Diese Arbeit wurde vollständig oder größtenteils in der Phase als Promovierender dieser Fakultät und Universität angefertigt;

2. Sofern irgendein Bestandteil dieser Dissertation zuvor für einen akademischen Abschluss oder eine andere Qualifikation an dieser oder einer anderen Institution verwendet wurde, wurde dies klar angezeigt;

3. Wenn immer andere eigene- oder Veröffentlichungen Dritter herangezogen wurden, wurden diese klar benannt;

4. Wenn aus anderen eigenen- oder Veröffentlichungen Dritter zitiert wurde, wurde stets die Quelle hierfür angegeben. Diese Dissertation ist vollständig meine eigene Arbeit, mit der Ausnahme solcher Zitate;

5. Alle wesentlichen Quellen von Unterstützung wurden benannt;

6. Wenn immer ein Teil dieser Dissertation auf der Zusammenarbeit mit anderen basiert, wurde von mir klar gekennzeichnet, was von anderen und was von mir selbst erarbeitet wurde;

7. Teile dieser Arbeit wurden zuvor veröffentlicht, siehe "List of Publications".

Datum

Unterschrift

Declaration on AI usage

I hereby declare that the use of artificial intelligence (AI) tools in the preparation of this document was strictly limited to improving grammar, punctuation, wording clarity, and overall flow of the text. No AI tools were utilized for generating original scientific content, data interpretation, or conceptual development. All research results, analyses, and scientific conclusions presented are entirely my own work.

List of publications

Parts of this thesis have already been published or presented at conferences. During my PhD studies, I co-authored the following publications:

Peer-reviewed publications

[1] ATLAS collaboration, *Measurement of the top-quark pole mass in dileptonic $t\bar{t}$ + 1-jet events at $\sqrt{s} = 13$ TeV with the ATLAS experiment*, *JHEP* **12** (2025) 023 [2507.02632]

[2] H.-Y. Bi, M. Kraus, M. Reinartz and M. Worek, *NLO QCD predictions for off-shell $t\bar{t}W$ production in association with a light jet at the LHC*, *JHEP* **09** (2023) 026 [2305.03802]

Preprint

[3] G. Bevilacqua, M. Reinartz and M. Worek, *SMEFT everywhere: a NLO study of $\boldsymbol{pp \to t\bar{t}H}$ with decaying tops*, 2605.11747

Conference proceedings

[4] M. Reinartz, *Off-shell $t\bar{t}W^+j$ production at NLO QCD accuracy*, *PoS* **ICHEP2024** (2025) 366 [2410.05959]

Relevant to this thesis is in particular Ref. [3].

Acknowledgements

I would like to thank Prof. Dr. Malgorzata Worek for her supervision and guidance during the many years I spent in her group. Additional thanks are extended to Prof. Dr. Robert Harlander, who is the second reviewer of this thesis. I want to further thank my collaborator Dr. Giuseppe Bevilacqua, who worked jointly on the project presented in this thesis and regularly lent me his expertise and advice throughout. I also would like to thank Prof. Dr. Michał Czakon for the fruitful discussions in the final phase of this project and Jonathan Hermann, who carried out initial explorations of the project. I extend special gratitude to Nikolaos Dimitrakopoulos, who shared an office with me for more than three years and was always available for conversations about both physics and life. I would like to further thank Leon Mans, Johannes Welter, Manal Alsairafi, Vilma Xhoxhaj, and Nikolaos Dimitrakopoulos for the many pleasant coffee breaks and for their help in proofreading this thesis. Finally, I want to thank my parents for their continuous support.

I also want to acknowledge that my work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft - DFG) under grant 396021762 - TRR 257: *Particle Physics Phenomenology after the Higgs Discovery.*

I also gratefully acknowledge the computing time provided to me at the NHR Center NHR4CES at RWTH Aachen University (project numbers `p0020216` and `rwth0414`). This is funded by the Federal Ministry of Research, Technology and Space, and the state governments participating on the basis of the resolutions of the GWK for national high performance computing at universities.

Contents

1 Introduction **1**

2 Theoretical Framework **15**
2.1 Theoretical predictions for hadron colliders . . . 15
2.1.1 Factorization and PDFs . . . 16
2.1.2 Hard-scattering processes . . . 18
2.1.3 Jet observables . . . 19
2.2 Intermediate resonances . . . 21
2.2.1 Off-shell calculation in the complex-mass scheme . . . 21
2.2.2 Narrow-width approximation . . . 23

3 Standard Model Effective Field Theory **27**
3.1 The Standard Model of Particle Physics . . . 27
3.2 Effective Field Theories . . . 31
3.3 Renormalization and operator mixing . . . 36
3.4 The SMEFT Lagrangian . . . 39
3.5 The Warsaw basis in the top-quark sector . . . 42

4 HELAC-NLO **47**
4.1 Born-level matrix element computations . . . 48
4.1.1 Recursive computation of helicity amplitudes . . . 48
4.1.2 Color-flow decomposition . . . 52
4.2 Virtual corrections . . . 54
4.2.1 Construction of 1-loop topologies . . . 55
4.2.2 Numerical computation of the numerator function . . . 56
4.2.3 Reduction of loop amplitudes . . . 58
4.2.4 The OPP method . . . 62
4.2.5 Reweighting techniques for virtual corrections . . . 64
4.3 Real-emission corrections . . . 66
4.3.1 Subtraction schemes . . . 66
4.3.2 Matrix elements in the soft limit . . . 67
4.3.3 Matrix elements in the collinear limit . . . 68
4.3.4 Catani-Seymour subtraction . . . 69
4.3.5 Nagy-Soper subtraction . . . 71
4.3.6 Subtraction for top-quark decays in the NWA . . . 73
4.4 Reweighting event files . . . 73

5 HELAC-SMEFT **79**
5.1 UFO interface . . . 82
5.1.1 Color decomposition . . . 83
5.1.2 Lorentz structure . . . 85
5.1.3 Technical description of the ECRU output files . . . 86
5.1.4 Code generation with HModelGenerator . . . 89
5.2 Implementing SMEFT in Helac-1Loop . . . 92
5.3 Optimization of the LO skeleton generation . . . 101
5.4 Implementing SMEFT in Helac-Dipoles . . . 104
5.5 Treatment of SMEFT event files in HEPlot . . . 105
5.6 Validation and benchmarks . . . 106
5.7 Computation of counterterm Feynman rules . . . 109

6 The $pp \to t\bar{t}H$ process in SMEFT **111**
6.1 Computational setup . . . 111
6.2 Integrated results . . . 118
6.3 Differential distributions . . . 127
6.4 Fixed and dynamical scale choice comparison . . . 137
6.5 Comparison with a stable top-quark computation . . . 141

7 Summary and Conclusions **149**

Chapter 1

Introduction

The discovery of the Higgs boson in 2012 by the ATLAS [5] and CMS [6] experiments at the Large Hadron Collider (LHC) has become the starting point for a large scale campaign to measure its fundamental properties. In the Standard Model (SM) many fundamental parameters are directly linked to the Higgs sector through the spontaneous symmetry breaking mechanism. As such, the Higgs sector itself is a powerful probe of the self-consistency of the theory. Over the past decade, many Higgs boson properties have been determined with high precision at the LHC. The combined result for the Higgs sector signal strength shows agreement with the SM description at the level of a few percent. More precisely, a 2022 combination of Run II data has found a signal strength ratio of $\mu = 1.002 \pm 0.057$ for CMS [7] and $\mu = 1.005 \pm 0.06$ [8] for ATLAS, where μ is constructed to compare several production and decay processes against their corresponding SM processes simultaneously. Similarly, the individual production and decay channels show great agreement with the SM within the current uncertainty estimates as can be seen in Figure 1.1, typically only differing by about one standard deviation when isolated to either production or decay. The uncertainties in the current experimental analysis are given almost equally by a statistical component, systematic uncertainties as well as theoretical modelling of the production and background processes. With Run III of the LHC planned to start in 2026, the experimental uncertainties are expected to further improve. Large emphasis is placed on the experimental determination of the Yukawa couplings. In the SM the coupling strength of Higgs boson interactions is directly related to the masses of the respective particles, which makes them a prime candidate to scrutinize the internal structure of the theory [9]. The size of the interaction increases with the particle masses. Consequently, the Yukawa interaction of the heaviest particles have been the first ones to be measured. Already using Run I data CMS and ATLAS were able to discover the $H \to \tau^+\tau^-$ decay channel [10–12] with 5.5 standard deviations. Initial evidence for the Higgs-bottom interaction was published in 2018 by CMS [13], with the full discovery being reported by both the ATLAS and CMS collaborations later in the same year [14, 15]. This was closely followed by the observation of the top-Higgs interaction [16–19]. Equivalent measurements for the first two fermion generations are significantly more difficult due to the smaller respective masses. The first experimental evidence for the Higgs-muon coupling was presented in Ref. [20], and more recently, constraints on the charm-quark Yukawa have also been obtained [21, 22].

The top-quark plays a special role in the SM. It is the heaviest of the fermions

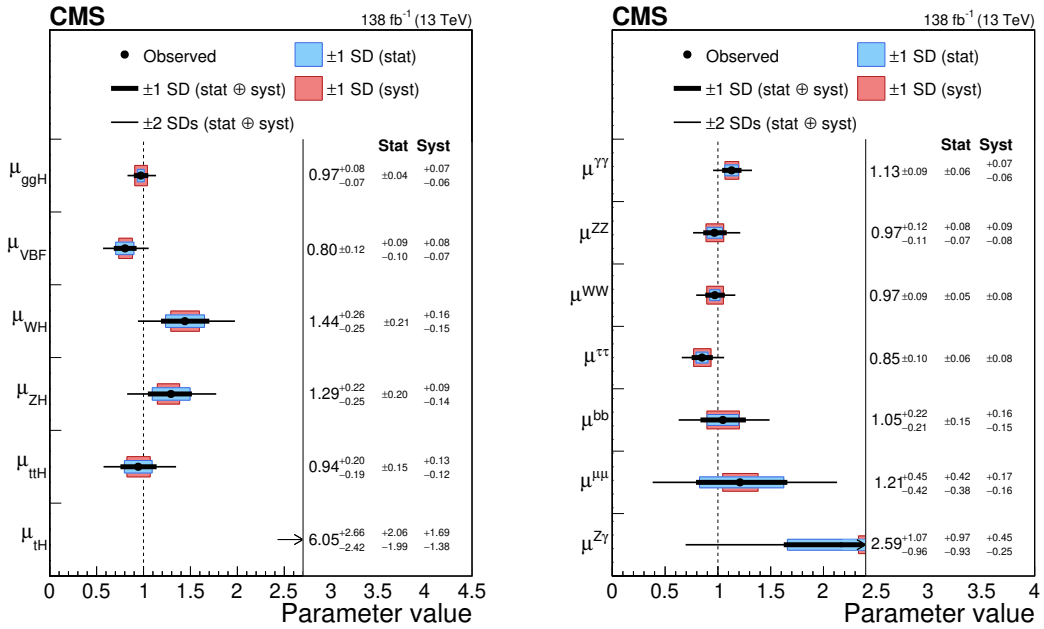


Figure 1.1: *Shown is the experimentally measured signal strength relative to the SM for various production processes (left) and decay channels (right). In order to disentangle the two contributions, the left plot assumes SM decay rates, while the right plot assumes that the production process is described by the SM. Figure taken from Ref. [7].*

with $m_t = 172.69 \pm 0.30\,\text{GeV}$ [23] and the only one with a Yukawa coupling close to unity $Y_t = \sqrt{2}\,m_t/v \approx 1$. The decay width of the top has been measured to be $\Gamma_t = 1.42^{+0.19}_{-0.15}\,\text{GeV}$ [23], which corresponds to a typical lifetime of $5 \times 10^{-25}\,\text{s}$. Because the QCD hadronization time scale exceeds the lifetime of the top, one only observes the top quark through its decay products. The most common decay channel is $t \to Wb$, which has branching ratio exceeding 95%. The W boson decays either leptonically into lepton and neutrino or hadronically into jets. In the context of $t\bar{t}$-associated productions, one commonly requires at least one leptonic decay to improve the clarity of the experimental signature. For the second decay channel, one must choose between the increased final state complexity of the additional jet activity or lower branching ratio of the leptonic decay chain. Furthermore, having two neutrinos in the final state increase the difficulty of the object reconstruction, because neither can be measured in the detectors at LHC. On the other hand, the semileptonic decay channel, where only one W-boson decays leptonically, offers a higher branching ratio and a direct correspondence between missing momentum and the neutrino. The top-quark decay products inherit the spin configuration from their mother particle during the decay, which makes the study of spin properties of top-quarks possible in LHC processes. Indeed, for many $t\bar{t}$-associated processes one can measure spin correlations between the top quarks through the angular distributions of the decay products. Interest in these observables is particularly pronounced in *new physics* searches, as many SM extensions modify the spin structure of SM vertices, leading to modified differential distributions for such spin-sensitive observables. Several SM phenomena can also be observed by investigating top-quark chirality. For example, quantum entanglement among top-quark pairs was recently observed at the LHC by carefully studying the aforementioned relative angular distributions [24].

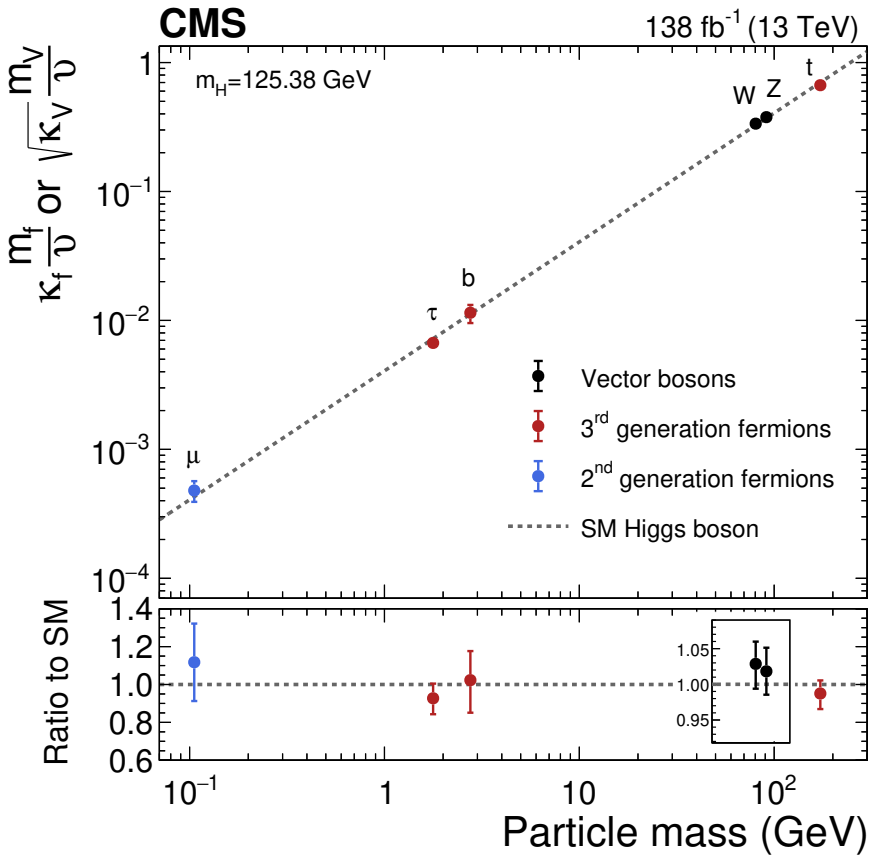


Figure 1.2: *Measured Higgs interaction strengths compared with the SM prediction (dotted line) at CMS using Run II data. Figure taken from Ref. [7].*

Because of the large top-quark Yukawa coupling one can often find overlapping interest between Higgs physics and top-quark physics at the LHC. Higgs production at the LHC comes in various processes. Representative diagrams are shown in Figure 1.3. The left panel of Figure 1.5 compares the production cross section across several center-of-mass energies. Despite being a purely loop-induced process in the SM the gluon-gluon fusion production rate is still the largest by far at LHC. It benefits from the large gluon contribution in the colliding protons and the large Yukawa coupling strength of the top-Higgs interaction. The next largest cross section corresponds to the vector boson fusion process $pp \to qqH$, which is suppressed by additional powers of the electroweak coupling constant. Similarly, the associated production with an electroweak vector boson sees a further reduction in the production rate. Finally, the Higgs production in association with two heavy quarks is even rarer, accounting for less than one percent of all produced Higgs bosons. The single-top production together with Higgs boson is also possible, but has not yet been observed. Instead a limit of about 6 times the Standard Model rate was established by CMS. Higgs and top physics interact closely in particular in the $pp \to t\bar{t}H$ and $pp \to H$ production modes. The connection to the top quark extends further to the Higgs decay modes. The right-hand side of Figure 1.5 compares the branching ratios of various Higgs decay channels. The decay into a photon pair $H \to \gamma\gamma$ is of particular significance due to the very clean experimental signature. Because the Higgs couples proportionally to the mass, no LO coupling to photons can occur. Instead, the interaction is me-

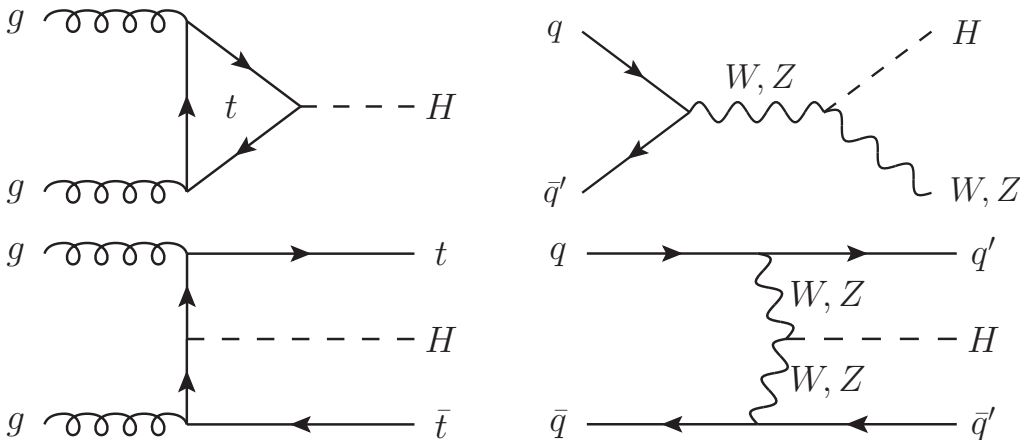


Figure 1.3: *Representative Feynman diagrams for Higgs production at LHC. Top-left is the gluon-gluon fusion, top right is the associated production with a vector boson, bottom-left is the $t\bar{t}H$ process and bottom-right is the production via vector-boson fusion.*

diated through a loop of top quarks or heavy vector bosons, which couple strongly to the Higgs. The decay through $ZZ^* \to 4\ell$ is similarly very sensitive, while other channels can suffer from the increased complexity of the hadronic jets generated in their decays.

The measurement of the Yukawa coupling can be performed using both direct and indirect methods. *Indirect methods* include the gluon-fusion process, where a Higgs boson is produced through a top-quark loop. Other approaches exploit the very large number of $t\bar{t}$ pairs available at the LHC to become sensitive to electroweak loop corrections in $pp \to t\bar{t}$ process [25–28]. The very rare four-top production process, $pp \to t\bar{t}t\bar{t}$, can also be used to constrain Y_t at tree-level already. Diagrams featuring one off-shell Higgs boson propagator can similarly be sensitive to Y_t [29–31]. These extractions strongly depend on the underlying model assumptions. In particular, the absence of small new physics interactions, which influence either top or Higgs physics is assumed in the extraction.

Direct searches provide a model independent way to measure the top Yukawa coupling. To that end, one measures final states involving both top-quarks and a Higgs boson. One such process is the single-top production in association with a Higgs boson, which can be used to study the charge-parity ($\mathcal{CP}$) properties of the Higgs interactions [33, 34].

Particularly important is the $pp \to t\bar{t}H$ process, which provides direct access to Y_t, albeit at the cost of a much lower production cross section than other Higgs production processes at the LHC. The aforementioned initial observation of the top-Higgs interaction was achieved within this process. To that end, evidence of the $H \to \gamma\gamma$, $H \to ZZ^* \to 4\ell$, $H \to$ multi-lepton and $H \to b\bar{b}$ decay channels were initially combined. A modern overview of the various channels is given in Figure 1.6. The cleanest experimental signature is observed for the $H \to \gamma\gamma$ decay process. However, the actual decay rate is small due to the channel being mediated by top-quark or W-boson loops. By contrast, the Higgs has much larger branching rates into massive states, most notably $H \to b\bar{b}$, but this channel comes at the cost of requiring the identification of b-jet flavors and has to compete with large QCD backgrounds

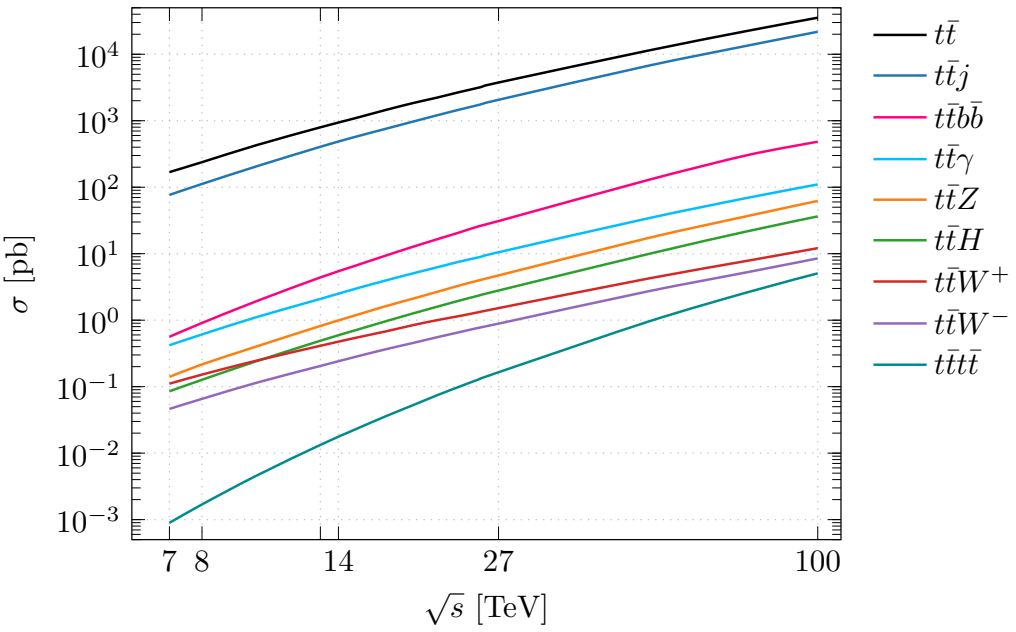


Figure 1.4: *Theoretical prediction for SM production cross-sections for $t\bar{t}$-associated processes as a function of the center-of-mass energy $\sqrt{s}$ at the LHC. Taken from Ref. [32].*

from $pp \to t\bar{t}b\bar{b}$ and $pp \to t\bar{t}jj$ production processes. Furthermore, the decayed final state involves four b-jets, which results in a combinatorial background for selecting the Higgs decay products. The modelling of the $pp \to t\bar{t}b\bar{b}$ background process itself is very challenging as well. As a result, the $pp \to t\bar{t}H$, $H \to b\bar{b}$ process is currently not in good agreement with the SM expectation. The treatment of Higgs decays is generally simpler than that of the top-quark decays, due to its scalar nature. Spin correlations do not have to be taken into account when incorporating these decays. Furthermore, the Higgs boson has a very small decay width with only a few MeV, which significantly increases the precision of approximations based on the ratio of the width over mass. Within such approximations the decay fully factorizes from the production process, which allows for separate treatment. Consequently, theoretical predictions often treat the Higgs as a stable particle and study the Higgs decay modes separately.

The $\mathcal{CP}$ nature of the top-Higgs interaction has also been studied at the LHC. While the $\mathcal{CP}$-even Higgs, as predicted in the SM, is compatible with the best fit results, one currently can not rule out an admixture of a $\mathcal{CP}$-odd component [37–42]. A graphical representation of the 2020 α_{CP} ATLAS measurement within gluon-gluon fusion production and $H \to \gamma\gamma$ decay channel is shown in Figure 1.7.

Theoretical predictions for the $pp \to t\bar{t}H$ process have seen large activity in the last two decades. Initial NLO QCD calculations in the stable top-quark approximation were performed by multiple groups [43–47] starting already in 2001. Further NLO EW corrections became available about a decade later in 2014 [48–50]. Complete-NLO corrections including all leading and subleading LO contributions alongside their NLO corrections to the stable process were computed in Ref. [51]. Further increases in perturbative accuracy are very challenging due to the three masses in the final state of the process. The flavor off-diagonal channels at NNLO

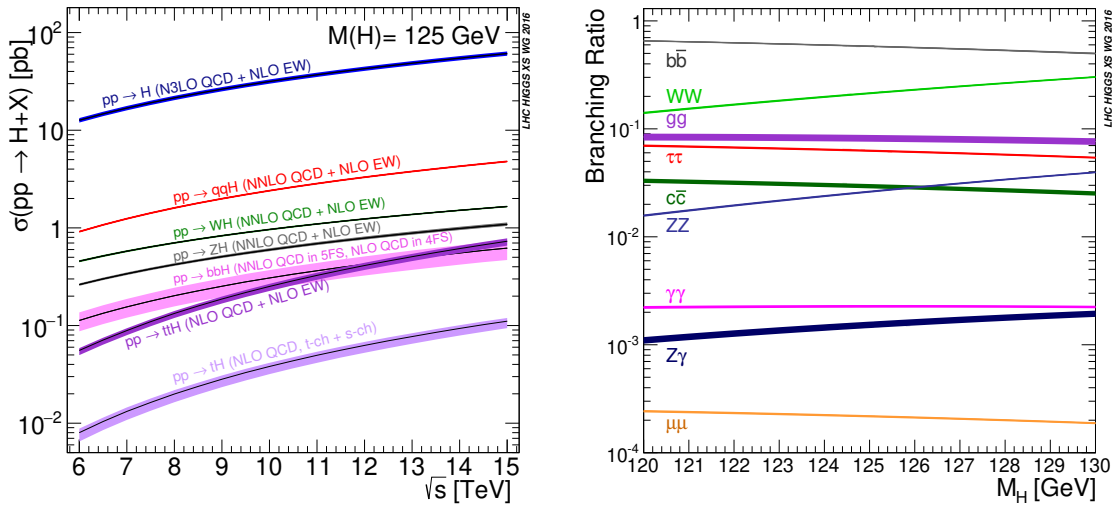


Figure 1.5: *Comparison of the cross section for several Higgs production processes at the LHC (left) and branching ratios of the Higgs decays (right). Both figures are taken from Ref. [35].*

were first computed in Ref. [52] and recently a full calculation including the diagonal subprocesses was presented in Ref. [53]. Within this calculation the two-loop matrix element was approximated using a soft Higgs approximation. Further improvements were achieved by combining the aforementioned approximation with a high-energy expansion in the small m_t limit [54]. The authors demonstrate that the two approximations yield consistent results and estimate the resulting uncertainty to be of $\mathcal{O}(1\%)$. Additionally, the results were combined with the complete NLO corrections. Despite the considerable difficulty of computing the exact two-loop correction, research on it has continued [55–57]. Beyond fixed order calculations, soft gluon contributions have been evaluated by multiple groups using resummation techniques up to next-to-next-to-leading-logarithmic precision (NNLL) and subsequently matched to the NLO predictions [58–61]. Colomb corrections have been similarly computed at NLL accuracy [62]. A recent state-of-art calculation combining the approximated NNLO results of Ref. [54] with NNLL soft-gluon resummation has been published [9]. The authors thoroughly discuss various sources of theoretical uncertainties stemming from the perturbative truncation, uncertainties in the PDFs, strong coupling constant as well as the top-quark mass m_t. For stable top-quark calculations, various differential distributions have been computed at NLO QCD accuracy and augmented with parton showers [63, 64]. In addition to the SM scalar, the former calculation also provides predictions for a pseudo-scalar Higgs. Recently, an approximate NNLO QCD computation has also been matched to parton showers [65] using the MiNNLOPS [66, 67] method. The code is publicly available within the POWHEG framework [68–71]. The simulation of top-quark decays has progressed significantly. Full off-shell calculations in the dileptonic decay channel, involving all double-resonant, single-resonant and non-resonant contribution as well as their interference effects and Breit-Wigner propagators for both the EW bosons and the top quark have been computed at NLO QCD in Ref. [72] and at NLO QCD and EW in Ref. [73]. These computations directly compute the full $pp \to \ell^+\nu_\ell\ell^-\bar{\nu}_\ell b\bar{b}H$ process. Higgs decays in the NWA have been additionally incorporated in the off-shell NLO

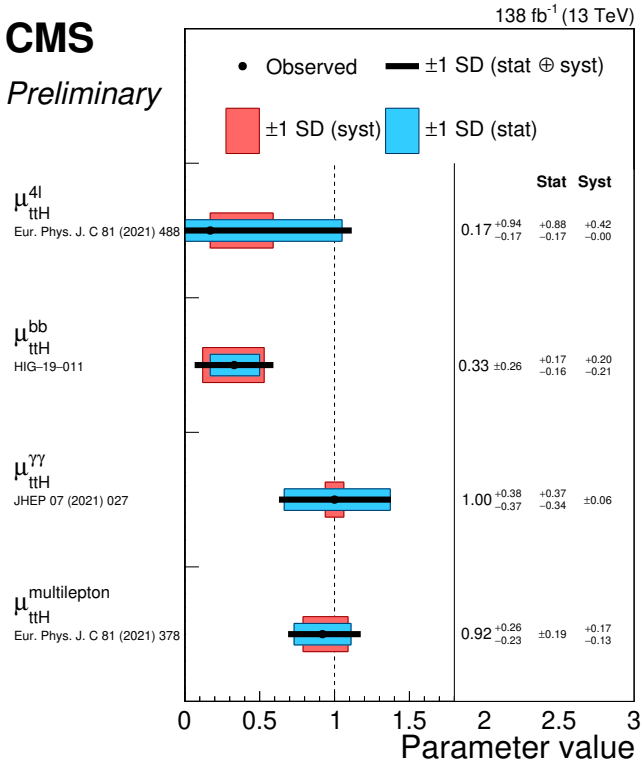


Figure 1.6: *Run II signal strength comparison in various decay channels of the* $pp \to t\bar{t}H$ *process at CMS. Taken from Ref. [36].*

QCD calculation in Ref. [74].

The SM has seen unprecedented success at the LHC, providing successful predictions across a large number of processes and observables. Nevertheless a complete description of nature necessarily must go *beyond the Standard Model* (BSM). The SM provides no satisfactory explanation for the matter-antimatter asymmetry in the universe and fails to provide a dark matter candidate. Another, more indirect, motivation can be found in the vacuum stability of the SM. The SM Higgs potential separately receives higher order corrections proportional to the top-quark Yukawa coupling Y_t as well as the self-interaction strength of the Higgs boson, which can be related back to the top-quark mass m_t and the Higgs mass m_H, respectively. The observed vacuum expectation value is only stable with respect to quantum tunneling effects if the minimum of the potential is indeed a global minimum. However, the loop corrections are opposite in sign and their respective magnitude thus decides the structure of the potential [75]. The various stability configurations in the m_t - m_H plane are shown in Figure 1.8. While current measurements point towards the SM vacuum being either stable or meta-stable, the SM does not actually provide an inherent mechanism to enforce this observation. The top and Higgs masses are both experimentally determined input parameters. Consequently, finding a more satisfactory explanation based on some underlying principle is highly desirable. The accidental relation between the top quark and Higgs boson provides further motivation to study the intersection of their respective regimes in LHC processes.

In light of the large success of the SM, constructing potential *new physics* theories as an extension to the SM seems like the most sensitive way forward. The lack of observed new resonances at the LHC indicates that possible new particles must

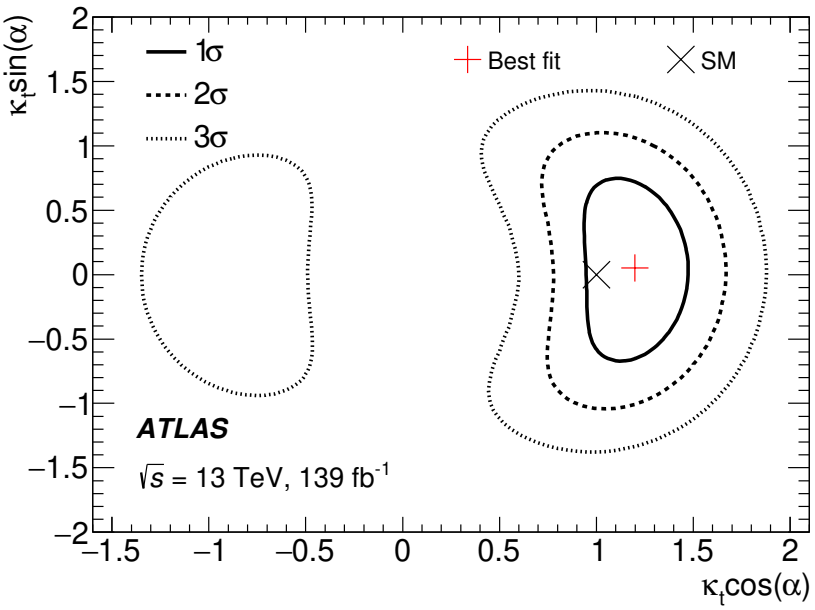


Figure 1.7: *Likelihood contours in the Higgs characterization model for the* $gg \to H \to \gamma\gamma$ *process measured by ATLAS.* $\sin(\alpha_{CP}) = 0$ *corresponds to a purely* $\mathcal{CP}$*-even interaction. Figure taken from Ref. [37].*

be either very weakly interacting or produced primarily at currently inaccessible energy ranges. Thus, many proposals predict new particle content at or above the TeV scale. Typically, they couple strongly to the Higgs sector and/or the top quark. Indeed, an admixture of a pseudoscalar component to the Higgs would be expected to first show up in the top Yukawa coupling due to the large top mass, which couples the two sectors. The $pp \to t\bar{t}H$ process is of particular interest for BSM physics. Deviations from the SM Yukawa coupling are likely to be observed here, as it provides a model-independent probe of the top-Higgs interaction. Indeed, predictions including $\mathcal{CP}$-odd Higgs admixtures within the process have been under investigation for some time [33, 34, 76, 77]. In these predictions the Lagrangian can be parametrised using the Higgs characterization model. Both $\mathcal{CP}$-even and $\mathcal{CP}$-odd states are incorporated and the mixture between them is described by an angle α_{CP}. Since many effective Lagrangians, such as the *two-Higgs-doublet model* (2HDM), can contain both components this formalism is an efficient way to study the potential impact of such extensions to the SM. It has been shown that approximations that work well within the SM, such as the narrow-width approximation (NWA), do not always translate well to BSM physics. For example, Ref. [77] shows that off-shell effects can become important even at the level of integrated cross-sections when considering a $\mathcal{CP}$-odd component in the top-Higgs interaction. To draw correct conclusions about potential BSM observations one must therefore carefully consider the accuracy of the modelling approaches used both in the SM predictions and the BSM extensions to it.

Another capable tool in the search for *new physics* is the *Standard Model Effective Field theory* (SMEFT). Within the SMEFT framework the SM is augmented with a series of non-renormalizable higher dimensional operators, which can capture potential effects of a more general theory at some high scale Λ in a largely model-

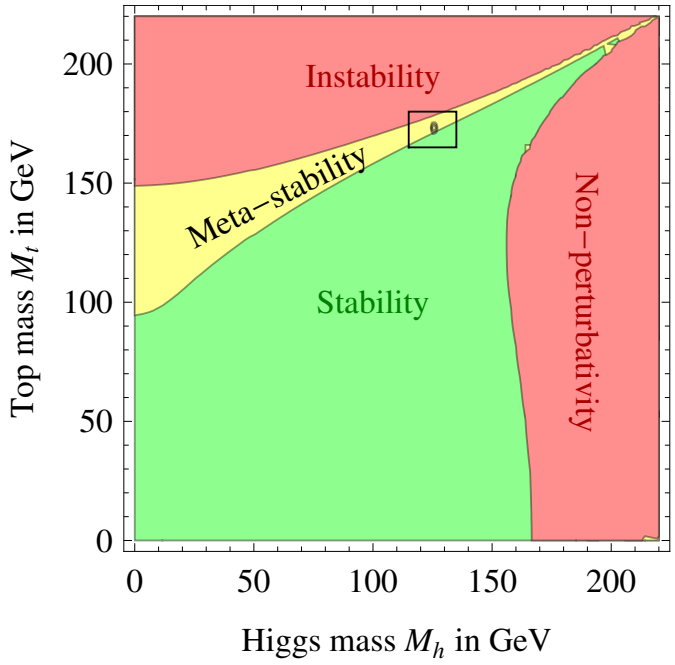


Figure 1.8: *Standard Model vacuum stability parametrised in the m_H - m_t plane computed up to two-loop precision. Figure taken from Ref. [75].*

independent way. At the much lower measurement scale one assumes that only the SM degrees of freedom are realized and that the SM gauge symmetry is fulfilled by all the additional effective operators. This flexibility is particularly useful in the absence of concrete hints towards specific BSM models. Many concrete BSM scenarios can be matched to SMEFT. Given the measured exclusion limits on SMEFT operators, one can then constrain the parameter space of the BSM model without performing dedicated experimental analyses. On the other hand, if specific operators within SMEFT can be measured with non-vanishing coefficients, then one can use this information as a basis in the search for a possible UV completion. Effective field theories are themselves predictive. Therefore, even if some BSM scenario can eventually be identified, one could still use existing SMEFT infrastructure as a powerful approximation apparatus. The study of SMEFT has attracted a lot of interest in recent years, because it is likely to yield a long-term returns in both experimental and theoretical fields alike.

SMEFT corrections appear in many sectors of LHC physics. A schematic overview of the contributions of paricular Wilson coefficients to several LHC datasets is presented in Figure 1.9. One can observe a large overlap between the Higgs and $t\bar{t}/t\bar{t}V$ measurements. The $pp \to t\bar{t}H$ process therefore seems particularly well suited to study new physics within the SMEFT framework.

Exclusion limits on the free parameters of the SMEFT have been extracted by both the CMS and ATLAS collaborations. The many processes involved in constraining new physics using the SMEFT prescription requires an equally large number of studies and measurement. A summary plot of the top quark - vector boson sector is shown in Figure 1.10. Shown are the 95% confidence level (CL) exclusions limits obtained from the combinations of several processes. To put the required experimental effort into context, we highlight that this summary plot was constructed by the top-quark EFT working group using Refs. [79–88]. The best

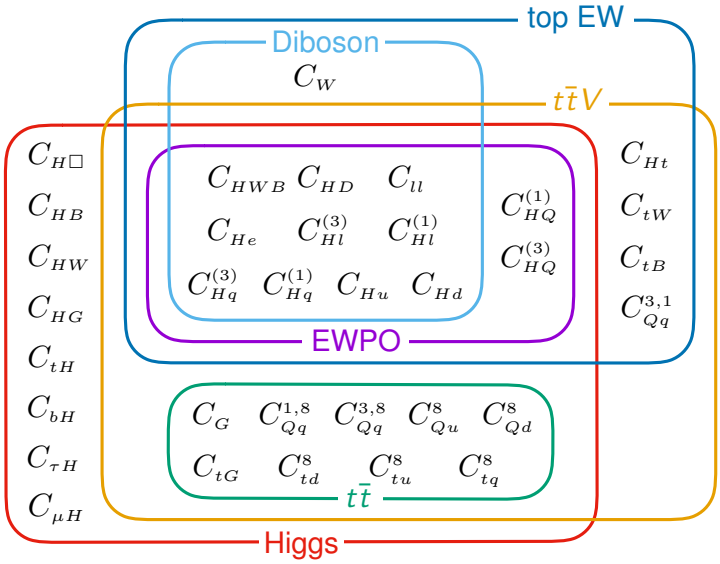


Figure 1.9: *Schematic dependence of SMEFT Wilson-coefficients on particular measurement sectors. Taken from Ref. [78].*

current exclusion limits in the top - vector-boson space range from roughly $|C_i/\Lambda^2| < 1\,\text{TeV}^{-2}$ to $|C_i/\Lambda^2| < 0.2\,\text{TeV}^{-2}$ in the most constrained cases. To contextualize these numbers, given an interaction strength $C_i \approx 1$, one would find $\Lambda \approx 2.2\,\text{TeV}$ for the latter case, where Λ is the characteristic mass scale of the underlying new physics. Similar summary plots for the scalar-top and four-fermion sector are available [89]. The SMEFT interactions between the top and Higgs boson, as well as four-fermion operators are typically somewhat less constrained with 95% exclusion limits at the level of $|C_i/\Lambda^2| < 10\,\text{TeV}^{-2}$ being quite typical. The most constrained operator in this category is restricted to about $|C_i/\Lambda^2| < 1\,\text{TeV}^{-2}$ [88]. A simultaneous fit of Wilson coefficients using various ATLAS measurements was recently performed in Ref. [90]. This study focused on Higgs-boson processes as well as $t\bar{t}$ events. The new physics scales corresponding to the $\mathcal{O}_{\phi G}$ operator are constrained to be above roughly 10 TeV at the 68% CL. The $\mathcal{O}_{t\phi}$ operator is less constrained. The authors quote an explored new physics range up to 500 GeV, while for $\mathcal{O}_{tG}$ BSM effects can be excluded below 1.8 TeV. The authors further find that including squared dimension-6 contributions can have an noticeable effect on the derived limits. However, overall no significant deviation from the SM was found.

Effective theories of this kind grow in importance with powers of the ratio E/Λ, where E is the energy scale of the observable. As a result, sensitivity to SMEFT effects can often be found in high-p_T regions of various processes. However, within top physics this regime is often a challenging environment, where off-shell effects, non-factorizable and electroweak corrections become important contributions. Significant theoretical effort to describe these regions in SM calculations has already been made. Naturally, theoretical descriptions of SMEFT are not as advanced, but a lot of progress has also been achieved already. Calculations in the top-quark sector have been performed in many processes, typically including key dimension-6 operators. Results are available for $pp \to t\bar{t}$ [91], single-top production [92–94] and various $t\bar{t}$-associated processes like $pp \to t\bar{t} + X$, with $X = \gamma, Z, H, t\bar{t}$ [91, 95–97]. Some of the aforementioned calculations include NLO QCD corrections, although

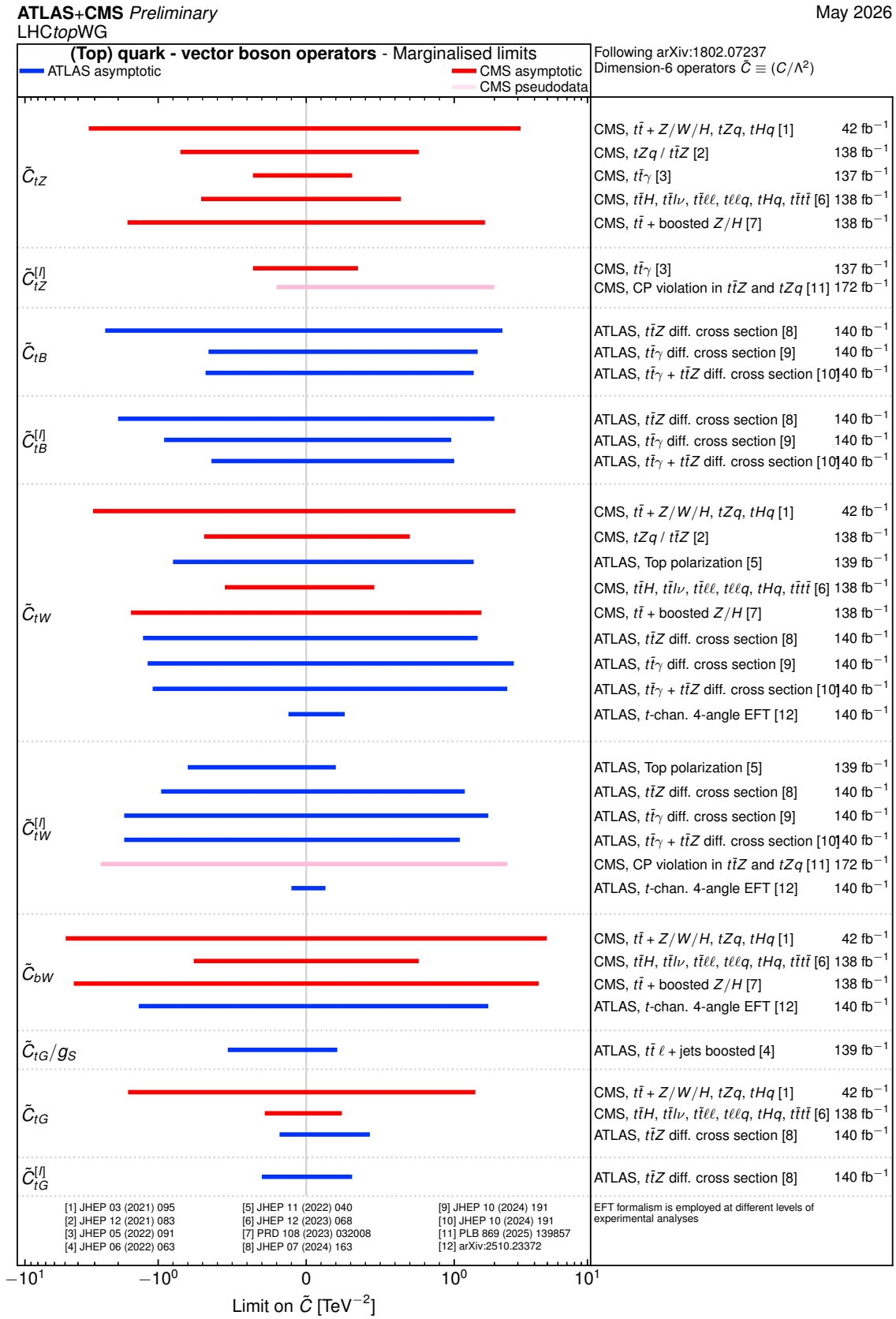


Figure 1.10: *Marginalised 95% confidence level exclusion limits on SMEFT Wilson coefficients containing the top quark and vector bosons using both CMS and ATLAS measurements. All Wilson coefficients contributing to a given process are treated as free parameters and limits obtained from different processes are combined as indicated in the right column. Taken from top EFT summary plots April 2026 available at [89].*

many are only performed at leading order. Global fits of the top-quark sector of the SMEFT have also been performed in Refs. [98, 99]. Many computer programs have been developed to facilitate these calculations. The Monte Carlo program MadGraph5_aMC@NLO [100] can support LO and NLO QCD calculation based on models supplied in the *Universal Feynman Output* (UFO) format [101, 102]. A LO implementation of SMEFT is available within the SmeftSim UFO model [103, 104]. At NLO QCD an implementation of the top-quark sector has been published in the Smeft@NLO UFO model [105]. This model contains some additional restrictions, but is sufficient for the calculation of $2 \to 3$ process at LHC for many important operators. Support for SMEFT computations has recently been added in a new major version of the Gosam3.0 program [106]. Feynman rules for SMEFT in Warsaw basis have been computed in Ref. [107] and made compatible with the FeynRules package [108, 109]. Furthermore, they have been implemented in the SmeftFR v3 Mathematica package [110], which supports Feynman rules up to mass dimension-8 consistently at LO. The renormalization group equations (RGEs) that connect Wilson coefficients at separate scales induce operator mixing within the SMEFT framework [111–113]. Recently, the mixing matrix for specific subsets of operators has been computed to two-loop order [114, 115]. Furthermore, the complete two-loop running for the baryon-number-conserving part of dimension-6 SMEFT was reported in Ref. [116]. The impact of running effects on SMEFT predictions for $pp \to t\bar{t}$ and $pp \to t\bar{t}H$ have been studied in Ref. [117] and Ref. [118], respectively. The effect of renormalization group effects on global SMEFT fits has been further investigated in Ref. [119]. In particular, the combination of several observables is needed to avoid the degeneracy induced through mixing effects, which otherwise limit the precision of derived bounds on the Wilson coefficients. Finally, the choice of γ_5 scheme is not arbitrary in SMEFT and only cancels when matched to an underlying UV theory. The impact of such scheme choices was scrutinized in Refs. [120, 121]. Computer codes to facilitate the running of SMEFT coefficients have been developed such as DsixTools [122, 123], Wilson [124] and RGESolver [125].

The study of SMEFT predictions so far is limited to the stable top-quark approximation or focuses exclusively on individual top-quark decays [126]. To extend predictions beyond this level, we compute the $pp \to t\bar{t}H$ process including the subsequent top-quark decays within the SMEFT framework. To that end, we develop a new Monte Carlo program, called Helac-Smeft, which is an overhauled and extended version of the original Helac-Nlo program [127], and can facilitate SMEFT calculations with unstable particle resonances in a consistent way. The new code focuses in particular on SMEFT calculation within the top-quark sector and aims to extend all existing functionality used for NLO SM computations to more general effective theories. However, NLO QCD calculations are supported for other processes outside the top-quark sector as well. SMEFT Feynman rules available in the literature are commonly implemented using the UFO format. The new program is built to interface to such results, which are already used in many other Monte Carlo codes. Most relevant to this thesis is the Smeft@Nlo UFO model.

This thesis is structured as follows. In Chapter 2 we discuss general aspects of fixed-order calculation for LHC predictions as well as the treatment of unstable particles. Chapter 3 contains a theoretical overview of effective field theory methods with particular focus on the SMEFT framework, which we adopt throughout this

thesis. Chapter 4 is dedicated to the methods used in the original HELAC-NLO program, which allows for the automated computation of LHC processes at NLO in QCD on the differential level. We discuss the calculation of both Born-level amplitudes, virtual corrections, real-emission contributions as well as the reweighting techniques used to efficiently perform the evaluation of theoretical uncertainties. The technical and conceptual modifications used to facilitate calculations in SMEFT within the program are discussed in Chapter 5. Finally, in Chapter 6 we discuss the results computed for the $pp \to t\bar{t}H$ process with top-quark decays included using the narrow-width approximation (NWA). Lastly, Chapter 7 summarizes the thesis and provides a conclusion.

Chapter 2

Theoretical Framework

In this chapter, we summarize important concepts used in the theoretical description of hadron-collisions at modern particle colliders like the Large Hadron Collider (LHC). We aim to establish the context in which the methods employed in this thesis are situated. To that end, we focus on the core pillars of event simulation as they are employed in many modern experimental and theoretical analyses. However, the vast activity in the field makes a comprehensive review of the topic unfeasible. The material presented in this chapter is largely based on Ref. [128], which we refer to for a more exhaustive overview of the field.

2.1 Theoretical predictions for hadron colliders

Scattering events at hadron colliders like the LHC pose a considerable challenge for accurate theoretical predictions. Due to QCD confinement, experiments cannot observe colored fundamental particles on their own. Instead, one observes uncolored composites. QCD is asymptotically free, that is the strong coupling constant α_s decreases as the energy scale increases. Mathematically the running of α_s with the energy scale can be written as a renormalization group equation given by

$$\mu^2 \frac{d\alpha_s(\mu^2)}{d\mu^2} = \beta(\alpha_s)\,, \tag{2.1}$$

where the function β can itself be expanded in powers of α_s as

$$-\beta(\alpha_s) = \sum_{n=0}^{\infty} b_n \alpha_s^{n+2} = \frac{\beta_0}{4\pi}\alpha_s^2 + \mathcal{O}(\alpha_s^3)\,. \tag{2.2}$$

For $SU(N)$ groups, one finds that the first coefficient is given by

$$\beta_0 = \frac{11}{3}C_A - \frac{4}{3}T_R n_f\,, \tag{2.3}$$

where the coefficients are given by $C_A = 3$ and $T_R = \frac{1}{2}$ for the $SU(3)$ group. The factor n_f is the number of active flavors involved in the running. The solution of Eq. (2.1) can be written in first approximation as [128]

$$\alpha_s(\mu^2) = \frac{1}{\frac{\beta_0}{4\pi}\log\left(\frac{\mu^2}{\Lambda^2_{\mathrm{QCD}}}\right)}\,. \tag{2.4}$$

The physics at hadron colliders spans energies from the TeV scale all the way down to the QCD confinement scale $\Lambda_{\text{QCD}} \approx 250\,\text{MeV}$. The description of the different sectors factorizes, which allows for separate treatment using different dedicated methods. At the highest scale, the partons inside the colliding hadrons can be treated as free fundamental particles. These partons partake in a scattering process. Because the QCD coupling is small for large energies, this regime is well suited to perturbative calculations. Scattering amplitudes are computed in a perturbative series, which is organized in the powers of the strong coupling constant and truncated at a *fixed-order*. Unstable particles produced in these processes further decay into leptons and quarks. Between the scales of the hard collision and the hadronization scale the particles further emit additional QCD and QED radiation, while continuously losing energy. This step can be modelled in parton showers, which can be matched to perturbative calculations to avoid double counting of emissions. At low scales the QCD coupling grows large and forces the colored partons to hadronize into colorless bound states. These composite states may either decay further or reach the detector directly, where they are measured as part of the signature of the collision event. Several Monte Carlo programs are available to simulate the various steps of the hadron collision events. In this thesis, we focus on the computation of fixed-order predictions. Wide-spread adoption of unfolding techniques allow experiments to compare with these predictions directly, which avoids the need to include all steps of the simulation chain in each theoretical analysis. Indeed, the physics of interest in high-energy particle physics is usually contained within the hard process. The precise description of hard scattering amplitudes is thus of particular importance.

2.1.1 Factorization and PDFs

The separation between the short-range physics of the hard process and the long-range physics associated with the structure of the proton is based on a factorization theorem. The cross section of the hadronic process with hadron momenta k_1 and k_2 is decomposed into a sum over sub-processes with partonic cross sections $\hat{\sigma}_{ab}$. The sums over a and b run over all partons in the hadron. For protons one typically includes gluons and quarks, but some calculations may also require the inclusion of photons or leptons. The latter are typically neglected due to their small contributions, but the photon contribution can be significant when computing electroweak corrections or when working with photon initiated processes. Two examples can be found in Refs. [129, 130]. Furthermore, the exact quark content can vary as some contributions can either be included inside the PDF definition or as part of the hard scattering process. The partons carry momenta $x_a k_1$ and $x_b k_2$ respectively and the contribution of a given parton momentum is weighted with the so called *parton-distribution-functions* (PDF), named here f_a and f_b. At LO the PDFs can be interpreted as the probability of finding a specific parton flavor in the proton with momentum x. At NLO this interpretation breaks down due to renormalization of the singularities, which appear in collinear QCD emissions. Overall the factorization can be written as

$$\sigma(k_1, k_2, \mu_R, \mu_F) = \sum_{a,b} \int_0^1 dx_a \int_0^1 dx_b \, f_a(x_a, \mu_F) \, f_b(x_b, \mu_F) \, \hat{\sigma}_{ab}(x_a k_1, x_b k_2; \mu_R, \mu_F) \tag{2.5}$$

where the $x_{a/b}$ is the momentum fraction carried by the two partons. The partonic process is computed at a reduced center-of-mass energy $\hat{s} = x_a x_b s$, where

$s = (k_1 + k_2)^2$ is the center-of-mass energy of the hadronic collision. The scale μ_F is the factorization scale at which the divergences of collinear QCD splitting is absorbed into the PDFs. A more thorough description of the singularity structure is given in Chapter 4. The evolution of the PDFs with the scale μ_F is given by the Dokshitser-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [131–134]. Explicitly, the DGLAP equation is given by

$$\mu_F^2 \frac{\partial f_a(x,\mu_F^2)}{\partial \mu_F^2} = \frac{\alpha_s}{2\pi} \sum_b \int_x^1 \frac{dz}{z} P_{ba}(z) f_b\left(\frac{x}{z}, \mu_F^2\right) \tag{2.6}$$

where P_{ij} are the Altarelli-Parisi splitting kernels. While the running of the PDFs can be computed perturbatively, the actual values can not be obtained this way. The PDFs are non-perturbative objects and are fitted from experimental data in practice. Extracting the PDFs from lattice calculations is, in principle, thinkable, but currently not competitive with the aforementioned data-driven approaches. Nevertheless, this direction remains a field of active research, due to the potential of computing the PDFs from first principles alone. Furthermore, experimental extractions are constrained by the underlying data, which can be less conclusive in specific regions of the phase-space. For a review, we refer to Ref. [135], and an example of recent study is given in Ref. [136].

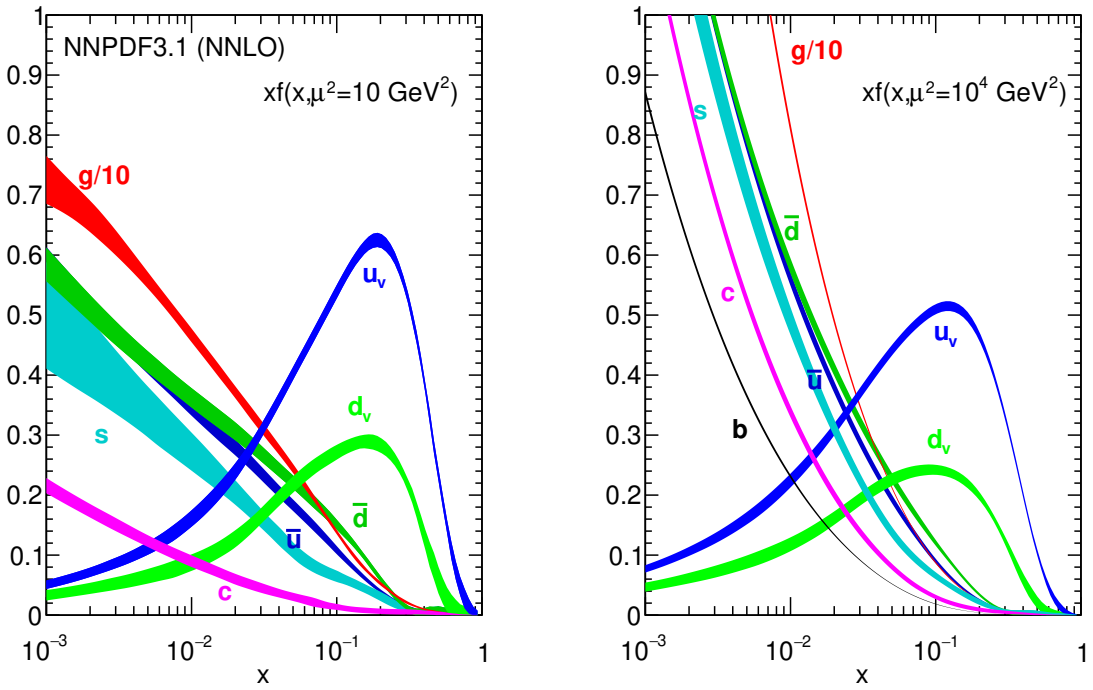


Figure 2.1: *Shown are quark and gluons PDFs in the NNPDF3.1 PDF set at NNLO with* $\mu_F^2 = 10\,\mathrm{GeV}^2$ *(left) and* $\mu_F^2 = 10^4\,\mathrm{GeV}^2$ *(right). Figure taken from Ref. [137].*

A key feature for practical cross-section calculations is that the PDFs constitute smoothly varying factors in the convolution in Eq. (2.5). They don't include any

narrow peaks, only an integrable increase as x tends 0. This makes them ideal for evaluations with Monte Carlo methods. Using conventional importance sampling techniques can yield almost constant weights with respect to the integration over the momentum fractions. The actual Monte Carlo variance is thus dominated by the underlying partonic cross section $\hat{\sigma}_{ab}$. We illustrate the point by showing the NNLO PDF set NNPDF3.1 [137] in Figure 2.1. The DGLAP evolution transfers probability from the valence-quarks to the sea-quarks and gluons when increasing the scale from $\mu^2 = 10\,\mathrm{GeV}^2$ to $\mu^2 = 10^4\,\mathrm{GeV}^2$, but the smooth behaviour remains intact.

2.1.2 Hard-scattering processes

The factorization formula in Eq. (2.5) expresses the collision process of two protons into a sum of partonic processes, which can be computed perturbatively. The task of fixed-order calculations is thus reduced to computing the partonic cross sections $\hat{\sigma}_{ab}$ with high accuracy before folding them with the respective PDFs. The partonic cross section can be written as

$$\hat{\sigma}_{ab} = \int d\Phi \frac{1}{S\,F} \left|\mathcal{M}_{ab}(\Phi, \mu_R, \mu_F)\right|^2 \mathcal{F}_J\,, \tag{2.7}$$

where $\mathcal{M}_{ab}$ is the scattering matrix-element of the process, S is the symmetry factor of the final state and F is the so-called flux factor given by

$$F = 4\sqrt{(p_a \cdot p_b)^2 - m_a^2\, m_b^2}\,, \tag{2.8}$$

where p_a and p_b are the momenta of the colliding partons. Furthermore, the integration measure for a process of partons with initial state momenta $p_{a/b}$ and n final state particles is given by

$$d\Phi(p_a, p_b; \{p_i\}) = \left(\prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3\, 2E_i}\right) (2\pi)^4 \delta^{(4)}\left(p_a + p_b - \sum_{i=1}^{n} p_i\right)\,, \tag{2.9}$$

where the energy of the i-th particle is given by $E_i = \sqrt{p_i^2 + m_i^2}$. Momentum conservation is implemented explicitly using a delta distribution. The integration itself is high-dimensional in practical calculations. For example, a process of $2 \to 7$ particles results in an integration over a 21-dimensional phase-space. Monte Carlo techniques scale well to high-dimensional integrations like this, but the adaptation of the phase-space optimization can be challenging for complicated integrand functions. The final state particles might be subject to additional kinematic constraints, which are implemented through the measurement function $\mathcal{F}_J$.

The calculation of the matrix elements is performed in a perturbative series. The amplitudes can be organized in terms of the powers of the QCD and EW couplings constants, denoted as α_s and α. The partonic cross section can then be written as

$$\hat{\sigma}_{ab} = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \left(\frac{\alpha_s}{4\pi}\right)^n \left(\frac{\alpha}{4\pi}\right)^m \hat{\sigma}_{ab}^{(n,m)}\,. \tag{2.10}$$

The electroweak coupling constant is much smaller than the strong one with values of $\alpha \approx \frac{1}{137}$ and $\alpha_s \approx 0.1$, respectively. The leading contribution, typically labelled

LO, in the expansion of Eq. (2.10) is typically the one with the lowest number of electroweak couplings as a result. Perturbative corrections can be organized in terms of the couplings as well. The NLO QCD corrections are given by the term with one additional power of α_s compared to the LO term and are usually the largest correction. Calculations up to NNLO QCD, i.e. up to α_s^2 relative to the leading term, should naively also include the first correction in α, because $\alpha \sim \alpha_s^2$ in terms of size. Furthermore, one can also include subleading contributions at Born-level. In this thesis, we only consider the LO and NLO QCD terms.

The evaluation of the matrix elements themselves can be achieved in various ways. Traditionally, one prepares a set of Feynman diagrams, which can be evaluated one by one. This approach is rather slow due to the large number of diagrams involved. Modern recursive algorithms can drastically reduce the required CPU resources by efficiently reusing common parts of different diagrams. We describe such an recursive algorithm in Section 4, which is based on the Dyson-Schwinger equations. For a more comprehensive overview, we refer the reader to Refs. [128, 138, 139].

2.1.3 Jet observables

Partonic cross sections computed at higher orders in perturbation theory contain several divergences. These cancel between the virtual corrections and real-emission contributions for *infrared- and collinear-safe* (IRC-safe) observables after UV renormalization and PDF factorization. One popular approach to constructing these IRC-safe quantities is via the use of jet algorithms, which can be implemented through the $\mathcal{F}_J$ function. In this section, we will briefly introduce the concept and present the most common algorithms.

A jet algorithm takes a set of final start partons in fixed-order calculations or a set of measured particles in experimental analyses and constructs a set of jets, which can be used to derive kinematical observables. Generally, such an algorithm should provide results that are robust under secondary emissions to allow the comparison of predictions and measurements. It turns out that this property also informs the finiteness of predictions in theoretical fixed-order calculations. While several such clustering algorithms have been developed and applied over the years, the most common ones used at LHC are based on the family of k_T-based algorithms. The general strategy is as follows [140]:

1. Create a list of pseudo-jets that contains all final state partons.
2. Compute the distances between all pseudo-jets i and j using a distance measure d_{ij}, as well as the distance to the beam-pipe given by d_{iB}.
3. If the smallest distance belongs to a pair i and j, replace them with a combined pseudo-jet ij. The momentum of the combined jet is given by the sum of their momenta $k_i + k_j$. Instead, if the smallest distance is of the d_{iB} type, move pseudo-jet i to the list of final-state jets.
4. Repeat the steps 2. and 3. until the list of pseudo-jets is fully exhausted.

The constructed jets can then be used to compute physical observables. The definition of the distance measure depends on the algorithm in question. The most

common algorithms for NLO calculations are the k_T [141], Cambridge/Aachen [142] and anti-k_T algorithms [140]. The distance measure of these is defined by

$$\begin{aligned} d_{ij} &= \min(k_{T,i}^{2l}, k_{T,j}^{2l}) \left(\frac{\Delta_{ij}}{R}\right)^2 , \\ d_{iB} &= k_{T,i}^{2l} , \end{aligned} \tag{2.11}$$

where $k_{T,i}$ is the transverse momentum of the i-th pseudo-jet and $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ with y_i and ϕ_i being the rapidity and azimuthal angle of pseudo-jet i, respectively. The integer l defines the algorithm in question: $l = 1$ is the k_T-algorithm, $l = 0$ the Cambridge/Aachen one and $l = -1$ the anti-k_T algorithm. The parameter R effectively defines the size of the jet radius. The exact kinematical features of the generated jets varies due to the chosen powers of the transverse momenta k_T. The k_T algorithm preferably clusters soft partons among themselves first, while the anti-k_T algorithm clusters soft partons to hard partons first. The latter strategy generally leads to more circular shapes of the generated jets, which is often deemed desirable in practice. For a more detailed comparison we refer to Ref. [140].

Lastly, we want to highlight the situation with identified flavors. In particular, separating b-flavored jets from other light jets is important in the context of top-quark physics, because tops predominantly decay through the $t \to Wb$ channel. When jet flavors are used to categorize observables, we must ensure that all singular contributions are grouped consistently by the jet flavor definition, such that the cancellation of divergences is preserved. In other words collinear and soft QCD splittings involving b-quarks are not allowed to change the derived jet flavors for a given phase-space point. At NLO QCD one implements this condition by combining jet flavors as

$$b\bar{b} \to g\,, \qquad bg \to b\,, \qquad \bar{b}g \to \bar{b}\,, \tag{2.12}$$

where the combinations bb and $\bar{b}\bar{b}$ are left as a scheme-dependent choice. The impact of the scheme-dependence was evaluated for the off-shell $pp \to t\bar{t}b\bar{b}$ process, i.e. with up to 5 b-flavored jets in the final state, in Ref. [143] and was generally found to be small compared to other theoretical uncertainties. At NNLO QCD the conditions outlined in Eq. (2.12) are no longer sufficient. To illustrate the point, consider a soft gluon splitting further into a soft $b\bar{b}$-pair as depicted in Figure 2.2. The two generated b-jets can be widely separated in angle and recombine with two other jets to pollute their respective flavors. Any observable that is exclusive with respect to the flavors of the two jets is not infrared-safe, i.e. it does not yield a finite result in perturbation theory. To combat this issue several approaches have been proposed. Extension to the traditional k_T-based algorithms have been developed, such as the flavor-k_T algorithm [144], which aim to resolve this issue, but usually sacrifice the consistent treatment of experimental data and theoretical calculations in this pursuit. Alternative approaches have also been proposed in Refs. [145, 146]. Furthermore, a comparative study has also been performed in Ref. [147].

This thesis focuses on calculations of fixed-order predictions at NLO in QCD. Therefore, the regular k_T-based algorithms are sufficient for our use-case and we refer the reader to the aforementioned references for the details on algorithms at NNLO and beyond.

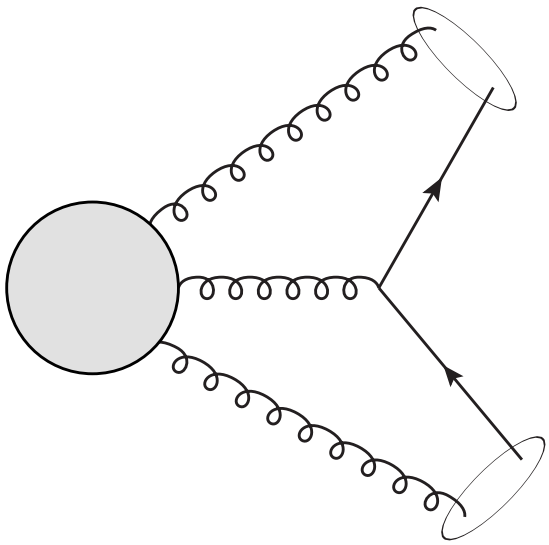


Figure 2.2: *Problematic configuration of flavored jets at NNLO. The core process is represented by the grey blob and emits two jets. A soft gluon splits into an additional pair of flavored quarks that polute the flavor of the two other jets. All Feynman diagrams in this thesis were created with the help of JaxoDraw [148].*

2.2 Intermediate resonances

Approximating top quarks and heavy bosons as stable particles is a well suited approach for the calculation of production rates with high precision. However, experiments can only measure fully decayed and hadronized final states. To bridge the gap between theory predictions and actual measurements, it is of considerable interest to treat intermediate resonances explicitly inside fixed-order calculations. In this section, we discuss two approaches that enable such studies. Full off-shell calculations using the complex-mass scheme (CMS) provide a comprehensive treatment, but come at the cost of large computational complexity. By contrast, the narrow-width approximation can describe particle decays consistently with low computational expense, although it neglects non-factorizable corrections to achieve this efficiency.

2.2.1 Off-shell calculation in the complex-mass scheme

First introduced in Ref. [149], the complex-mass scheme provides a gauge invariant way to include intermediate resonances in perturbation theory. The scheme has since been expanded to NLO corrections [150, 151]. The CMS works by introducing a set of complex renormalization constants. Bare masses of unstable particles are written as [152]

$$m_B^2 = \mu^2 + \delta\mu^2 \,, \tag{2.13}$$

where m_B is the bare mass, μ^2 is the renormalized complex-mass and $\delta\mu^2$ is a counterterm. The complex mass is resummed into the propagator expression, while $\delta\mu^2$ generates a counterterm vertex that is included in higher order calculations. The propagator is given by the familiar expression

$$\Delta(p,\mu) = \frac{1}{p^2 - \mu^2} \,, \tag{2.14}$$

where the pole-prescription is defined by the imaginary part of μ^2. The value of the complex mass is set to the location of the complex pole of the amplitude, that is

$$\mu^2 = m^2 - im\Gamma\,, \tag{2.15}$$

where m is the real renormalized mass and Γ is the decay width of the resonance. We note that this procedure is simply a reordering of the Lagrangian. No term has been added or removed. In practical implementations, one can simply replace the renormalized real masses with the complex masses and compute derived quantities like the weak mixing angle and Yukawa interaction strength in terms of the complex masses. The renormalization conditions are then adjusted similarly, e.g. for the on-shell scheme one requires

$$\Sigma_R(p^2)\Big|_{p^2=\mu^2} = 0\,, \qquad (\tfrac{d}{dp^2}\Sigma_R)\Big|_{p^2=\mu^2} = 0\,, \tag{2.16}$$

where Σ_R is the self-energy after renormalization. This condition fixes the mass and wavefunction renormalization constants by requiring that external states correspond to the physical particle states. Real-valued bare Lagrangians further imply

$$\mathcal{I}(\mu^2) = -\mathcal{I}(\delta\mu^2)\,, \tag{2.17}$$

which can be used to test the consistency of the derived expressions. The location of the complex pole is gauge independent, which ensures gauge invariant predictions assuming that the renormalization conditions are gauge independent. Within the CMS, the invariant-mass distributions of resonant contributions are described by the Breit-Wigner distributions, which can be written as

$$\mathcal{B}(p^2) = \frac{1}{(p^2 - m^2)^2 + m^2\Gamma^2}\,. \tag{2.18}$$

Full off-shell calculations specify the exact initial- and final-state of the fully-decayed process and consistently include all Feynman diagrams that appear at a given order in perturbation theory. For example, the $pp \to t\bar{t}H$ process with top quarks decaying in the dileptonic channel is specified by

$$pp \to \nu_e e^+ \bar{\nu}_\mu \mu^+ b\bar{b}H\,. \tag{2.19}$$

Representative Feynman diagrams for this process are shown in Figure 2.3. The left diagram is a double-resonant contribution, the middle one is single-resonant and the right one is non-resonant. The full off-shell approach incorporates all of these diagrams as well as all their interference effects and treats the intermediate unstable particles, such as top quark, W boson and Z boson, using the CMS. Hybrid schemes are also valid. For example, the top quarks can be included using the CMS, while the electroweak vector bosons could be treated using the NWA instead. While the off-shell approach provides the most complete description of unstable particles for fixed-order calculations, the increase in number of diagrams significantly increases the cost of the amplitude evaluation. In fact, while the inclusion of all single- and non-resonant diagrams for multiple unstable particles often massively increases the required CPU resources, one will typically not find very large corrections on the level of integrated cross sections in the SM. Furthermore, the reduction of 1-loop amplitudes with additional internal propagators is often required, which comes with its own additional cost.

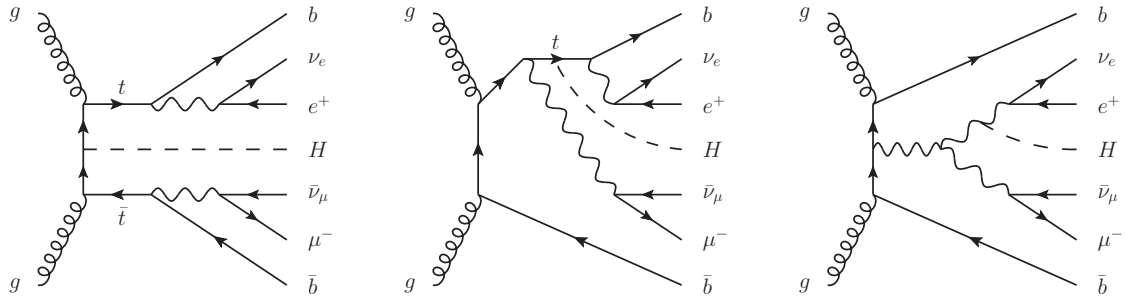


Figure 2.3: *Representative diagrams for the $pp \to t\bar{t}H$ process in the di-leptonic decay channel. The left diagram contains two top resonances, the middle one just one top resonance and the right one includes no top resonance.*

2.2.2 Narrow-width approximation

The narrow-width approximation is an approximation in which one treats unstable particles by taking the limit $\frac{\Gamma}{m} \to 0$. This approach leads to several simplifications that reduce the required computational cost of the calculation. In this approximation the Breit-Wigner distribution simplifies to

$$\lim_{\frac{\Gamma}{m}\to 0} \frac{1}{(p^2 - m^2) + m^2\Gamma^2} = \frac{\pi}{m\Gamma}\delta(p^2 - m^2) + \mathcal{O}\left(\frac{\Gamma}{m}\right), \tag{2.20}$$

where the δ-function forces the intermediate resonances on-shell. Consequently, the amplitude only contains contributions that exactly contain the intermediate resonances and single- or non-resonant contributions are no longer included. The process can be written as a production process and a series of subsequent decay processes. The computational gain is two-fold. First, the number of diagrams is reduced due to the omission of non-factorizable terms. Second, the loop-amplitudes no longer include corrections that exchange momenta between the production and decay processes, as these would take the intermediate unstable particles off-shell. Generally, this reduces the number of propagators in loops, which can strongly reduce the cost of their respective evaluation as will be described in Section 4.2. For example, the $pp \to \nu_e e^+ \bar{\nu}_\mu \mu^+ b\bar{b}H$ process will contain heptagons, when computed in the off-shell approach, but can contain at most pentagons in the NWA. The computation of one-loop correction in modern NLO Monte Carlo tools typically requires repeated evaluation of the numerator for any given loop structure, the number of which increases with the complexity of the underlying topology. Thus, any simplification at this stage comes with a huge improvement to the overall runtime. The phase-space of the final-state particles is simplified by the additional constrains induced by the on-shell conditions of the intermediate particles. This leads to better optimization of the phase-space and faster Monte Carlo integration as a result. In the context of modern Monte Carlo programs, one can expect one additional improvement for the SM specifically. W-fermion couplings are chiral, which restrict the helicity states that can contribute to an amplitude that involves them. In the context of $pp \to t\bar{t}H \to W^+W^-b\bar{b} \to \nu_\ell \ell^+ \bar{\nu}_\ell \ell^+ b\bar{b}H$, one can conclude that the helicity configuration of the final state is fully restricted to only one possibility in the NWA. Overall, one thus only has to compute four helicity states, which originate from the different initial-state assignments. This represents a massive reduction

from the naively expected $2^8 = 256$ helicity states that one could expect. In the off-shell case, this restriction can be partially lifted by non-resonant contributions. The use of Monte Carlo techniques, described in Chapter 4, can partially offset this additional cost. However, even in that case, one observes less efficient sampling of helicity configurations for the off-shell calculation compared to the NWA.

The difference to the off-shell calculation is naively expected to be of $\mathcal{O}\left(\frac{\Gamma}{m}\right)$. For the top quark one finds $\Gamma_t/m_t \approx 0.8\%$, for the W-boson $\Gamma_W/m_W \approx 2.6\%$ and for the Z-boson $\Gamma_Z/m_Z \approx 2.8\%$. In practice, this expectation only holds for inclusive observables. Observables can become very sensitive to these corrections in kinematical regimes, which isolate the phase space of specific resonant contributions. Consider the example of $M_{b\ell}^{\min} = \min\left(M_{b_1\ell}, M_{b_2\ell}\right)$ in the context of $t\bar{t}$-associated processes in the dileptonic decay channel. Double-resonant contributions are limited here to the phase-space region below $\sqrt{m_t^2 - m_W^2} \approx 155\,\text{GeV}$. Starting from this edge, one can expect very large off-shell corrections, which allows experiments to probe the off-shell nature of the top quark. A representative example is shown for the $pp \to e^+\nu_e\mu^-\bar{\nu}_\mu b\bar{b}H + X$ process in Figure 2.4. The figure compares the SM predictions with two possible admixtures of $\mathcal{CP}$-odd top-Higgs interactions. One can further observe that off-shell corrections are enhanced within the BSM scenarios.

The NWA matches current experimental analyses more closely than full off-shell calculations. Indeed, typical setups divide events into signal and background processes. The production processes can then be defined as the double-resonant contribution and a single-resonant contribution as a background process. This approach does induce an additional uncertainty and is conceptionally problematic, as interference effects can not be uniquely classified like this. Comparing with off-shell predictions can reduce such uncertainties. Furthermore, the validation of the SM requires that such effects are measured. In the context of BSM studies, one often investigates high-p_T tails or distributions with kinematical edges, which are both sensitive to potential BSM effects and to off-shell effects alike. Consequently, off-shell calculations play an important role in correctly interpreting results in these phase-space regions.

Returning to the $pp \to t\bar{t}H \to W^+W^-b\bar{b} \to \nu_e e^+ \bar{\nu}_\mu \mu^+ b\bar{b}H$ process, the differential cross section in the NWA factorizes into a production rate and branching ratios for the subsequent decays like

$$\begin{aligned} d\sigma_{\text{NWA}}^{\text{LO}} &= d\sigma_{t\bar{t}H}^{\text{LO}} \frac{d\Gamma^0_{t\to W^+b} d\Gamma^0_{\bar{t}\to W^-\bar{b}}}{\left(\Gamma_t^{\text{LO}}\right)^2}\,, \\ d\sigma_{\text{NWA}}^{\text{NLO}} &= d\sigma_{t\bar{t}H}^{\text{NLO}} \frac{d\Gamma^0_{t\to W^+b} d\Gamma^0_{\bar{t}\to W^-\bar{b}}}{\left(\Gamma_t^{\text{NLO}}\right)^2} \\ &\quad + d\sigma_{t\bar{t}H}^{\text{LO}} \left(\frac{d\Gamma^1_{t\to W^+b} d\Gamma^0_{\bar{t}\to W^-\bar{b}}}{\left(\Gamma_t^{\text{NLO}}\right)^2} + \frac{d\Gamma^0_{t\to W^+b} d\Gamma^1_{\bar{t}\to W^-\bar{b}}}{\left(\Gamma_t^{\text{NLO}}\right)^2} \right)\,, \end{aligned} \tag{2.21}$$

where the LO and NLO production rates are given by $d\sigma_{t\bar{t}H}^{\text{LO}}$ and $d\sigma_{t\bar{t}H}^{\text{NLO}}$, respectively. The LO and NLO differential decay rates are denoted by $d\Gamma^0_{\bar{t}\to W^-\bar{b}}$ and $d\Gamma^1_{\bar{t}\to W^-\bar{b}}$. The cross section is then (N)LO accurate when the decay width of the top quark is computed at (N)LO in QCD. The decays of the W bosons into lepton and neutrino are suppressed here for simplicity, but are also taken in to account using the NWA. At NLO QCD they always enter at LO, but some NLO EW effects can be incorporated by setting the W-boson width to its NLO value. Furthermore, we suppress

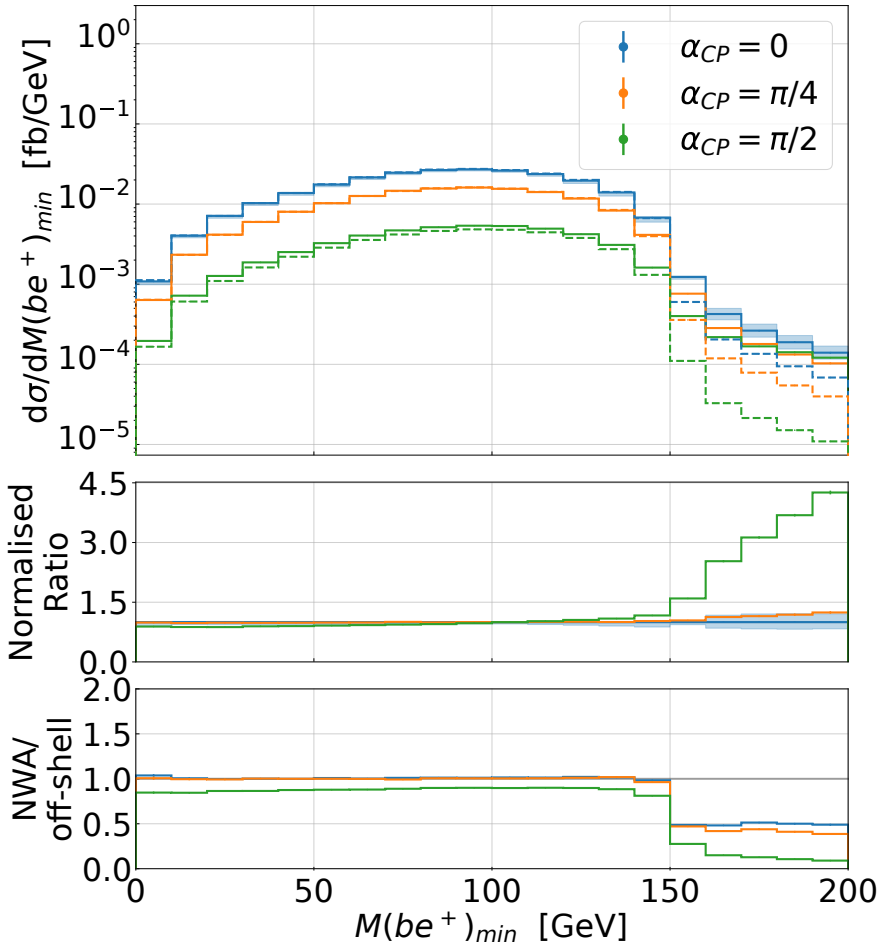


Figure 2.4: *Comparison of the $M_{b\ell}^{min}$ distribution in $pp \to e^+\nu_e\mu^-\bar{\nu}_\mu b\bar{b}H + X$ production within the Higgs characterization model. The angle α_{CP} quantifies the size of the $\mathcal{CP}$-odd component. The last panel shows the ratio of the off-shell calculation to the NWA. A sharp drop-off can be observed after the kinematical edge at roughly 155 GeV. Results were evaluated using $\mu_R = \mu_F = H_T/2$, where H_T is the scalar sum of the final-state transverse momenta and a center-of-mass energy of $\sqrt{s} = 13\,\text{GeV}$. The PDF set is set to NNPDF3.1 with $\alpha_s(m_Z) = 0.118$. Figure was taken from Ref. [77].*

helicity indices, but note that spin correlation between the production process and the decaying top quarks are fully taken into account within this approach. The expression in Eq. (2.21) still contains spurious higher order terms in the α_s expansion, because the NLO top-quark width itself depends on the strong coupling constant. Consequently, one should re-expand Eq. (2.21) in powers of α_s, which leads to

$$d\sigma^{\text{NLO}}_{\text{NWA,exp.}} = d\sigma^{\text{NLO}}_{\text{NWA}} \times \left(\frac{\Gamma_t^{\text{NLO}}}{\Gamma_t^{\text{LO}}}\right)^2 - d\sigma^{\text{LO}}_{\text{NWA}} \times 2\frac{\Gamma_t^{\text{NLO}} - \Gamma_t^{\text{LO}}}{\Gamma_t^{\text{LO}}}\,. \tag{2.22}$$

Results in the NWA in this thesis are always calculated based on Eq. (2.22), but we note that large deviation between the expanded and unexpanded results often point to poor perturbative convergence. For well behaved computations one thus should not expect large differences.

Chapter 3

Standard Model Effective Field Theory

This thesis is primarily concerned with the calculations in the Standard Model Effective Field Theory framework. In this chapter, we briefly review the structure of the Standard Model of particle physics and introduce the notation used in the rest of the thesis. We then discuss Effective Field theories generally and finally focus on the case of SMEFT in particular.

3.1 The Standard Model of Particle Physics

The Standard Model of Particle Physics is a gauge theory defined by the $SU(3) \times SU(2) \times U(1)$ gauge group. Gauge theories are invariant under local gauge transformation, that is the Lagrangian density $\mathcal{L}$ stays intact under any transformation of the type $\psi \to U(x)\,\psi$, where ψ is a vector of fields transforming under the group and $U(x) = e^{i\omega_a(x)t^a}$ is a group member of the gauge transformation. The parameters $\omega_a(x)$ are spacetime dependent functions and t^a are the generators of the respective group. Derivatives of transformed fields generate additional $\partial_\mu\,\omega_a(x)$ terms, which spoil the invariance of the Lagrangian under the gauge transformation. To cure this behavior one introduces the covariant derivate through the replacement

$$\partial_\mu \to D_\mu = \partial_\mu - i\,g\,t^a A^a_\mu\,, \tag{3.1}$$

where g is the coupling strength of the theory and A^a_μ are the gauge fields. The gauge fields transform under the gauge transformation so that

$$D_\mu\psi \to U\,(D_\mu\psi)\,. \tag{3.2}$$

In other words, one requires that the covariant derivative of a field transforms in the same way as the field itself. This restriction leads to the transformation property of the gauge field, which can then be written as

$$A^a_\mu t^a \to U A^a_\mu t^a U^{-1} - \frac{i}{g}\,(\partial_\mu U)\,U^{-1}\,. \tag{3.3}$$

The replacement of $\partial_\mu \to D_\mu$ introduces an additional interaction term between the fields ψ and the gauge boson. The kinetic term of the gauge field is given by

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4}F^a_{\mu\nu}\,F^{a\mu\nu}\,, \tag{3.4}$$

where the field-strength tensor $F_{\mu\nu} = t^a F^a_{\mu\nu}$ can be defined as

$$F_{\mu\nu} = \frac{i}{g} \left[D_\mu, D_\nu \right] . \tag{3.5}$$

The field-strength tensor transforms under the gauge transformation as $F_{\mu\nu} \rightarrow U(x) F_{\mu\nu} U(x)^\dagger$. The gauge fields are themselves charged in non-abelian gauge groups like $SU(3)$ and $SU(2)$. However, for abelian groups like $U(1)$, the corresponding gauge boson is not charged and thus does not couple to itself.
The field content of the SM can be categorized into fermions and bosons. The SM contains gauge bosons for each of the three groups: The gluon G^a_μ, which mediates Quantum Chromodynamics (QCD), corresponds to the $SU(3)$ group. The $SU(2)$ interaction is mediated through the left-handed W-bosons W^I_μ, and the $U(1)$ gauge field is denoted by B_μ. The coupling constant of the strong interaction is called g_s, while the $SU(2)$ and $U(1)$ couplings are given by g and g', respectively. The charges under $SU(3)$ are called color, while the ones under $SU(2)$ are characterized by weak isospin I_W and the third component I_{W_3} as is customary for angular momentum quantum numbers. The $U(1)$ charge is called hypercharge Y and is normalized with an additional factor of two. The hypercharge can be expressed as

$$Y = 2(Q_0 - I_{W_3}) , \tag{3.6}$$

where Q_0 is the charge of the field in question. The covariant derivative can be written as

$$D_\mu = \partial_\mu - i g_s t^a_{R_3} G^a_\mu - i g t^I_{R_2} W^I_\mu - i g' t_{R_1} B_\mu , \tag{3.7}$$

where the generator operators $t^a_{R_3}$, $t^I_{R_2}$ and t_{R_1} denote the group generators in the representation R_3, R_2 and R_1 of the field on which D_μ acts. In the fundamental representation, the $SU(3)$ generators are given by the Gell-Mann matrices through $\frac{\lambda^a}{2}$ and in the adjoint representation by the structure constants of the group $-if^{abc}$. For the $SU(2)$ group one instead finds the Pauli-matrices $\frac{\sigma_I}{2}$ and the Levi-Civita tensor $-i\epsilon^{IJK}$ for the fundamental and adjoint representations, respectively.

The fermions transforming under $SU(3)$ are called quarks. They can either be left-handed and transform in the fundamental representation of $SU(2)$, or right-handed, which leaves them invariant under $SU(2)$ gauge transformations. The left-handed quarks are organized into three generations of quark doublets Q_i given by

$$\begin{pmatrix} u \\ d \end{pmatrix} , \qquad \begin{pmatrix} c \\ s \end{pmatrix} , \qquad \begin{pmatrix} t \\ b \end{pmatrix} . \tag{3.8}$$

The up-type quarks are up u, charm c and top t, while the down-type quarks are down d, strange s and bottom b. The right-handed variants can be organized into three-tuples u_i and d_i instead. Up-type quarks carry electric charge $+\frac{2}{3}$, while down-type quarks carry $-\frac{1}{3}$. In this thesis, we will often be particularly interested in the third generation. For brevity we will then always refer to the third left-handed doublet as Q with no index and the right-handed top-quark as t. The QCD Lagrangian can be written as

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} + i \sum_{i=1}^{3} \bar{Q}^{(k)}_i \not{D}_{kl} Q^{(l)}_i + i \sum_{i=1}^{3} \bar{u}^{(k)}_i \not{D}_{kl} u^{(l)}_i + i \sum_{i=1}^{3} \bar{d}^{(k)}_i \not{D}_{kl} d^{(l)}_i , \tag{3.9}$$

where we have made the color indices k, l explicit. Similarly, the leptons can be organized into three generations as well. The lepton sector is not charged under $SU(3)$, but can couple to $SU(2) \times U(1)$. We will refer to the left-handed lepton doublets as L_i and the right-handed lepton fields as e_i. The left-handed leptons are organized into $SU(2)$ doublets as

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}, \qquad \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}, \qquad \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}. \tag{3.10}$$

The neutrinos don't carry electromagnetic charge, while the other leptons carry charge -1. The Lagrangian of the electroweak sector can be written as

$$\mathcal{L}_{\text{EW}} = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} \vec{W}_{\mu\nu} \cdot \vec{W}^{\mu\nu} + i \sum_{i=1}^{3} \bar{L}_i \not{D} L_i + i \sum_{i=1}^{3} \bar{e}_i \not{D} e_i \,. \tag{3.11}$$

Because neutrino masses are negligible for LHC observables, we adopt the minimal SM field content and omit right-handed neutrinos ν_i (which are gauge singlets and have no renormalizable interactions), so no kinetic or Yukawa term for ν_i appears in the electroweak Lagrangian. A gauge-fixing term is required to define the propagators of the gauge bosons. In the R_ξ gauge one choses

$$\mathcal{L}_{GF} = -\frac{1}{2\xi_G} (\partial^\mu G^a_\mu)^2 - \frac{1}{2\xi_W} (\partial^\mu W^I_\mu)^2 - \frac{1}{2\xi_B} (\partial^\mu B_\mu)^2 \,, \tag{3.12}$$

and supplements it with the Faddeev-Popov ghost Lagrangian to remove unphysical longitudinal degrees of freedom inside loops. In this thesis, we only compute NLO QCD corrections and therefore only have to supply the corresponding ghosts for QCD. The ghost Lagrangian is given by

$$\mathcal{L}_{\text{Ghost}} = -\bar{c}^a \left(\partial^\mu D^{ab}_\mu \right) c^b \,, \qquad D^{ab}_\mu = \delta^{ab} \partial_\mu - g_s f^{abc} G^c_\mu \,, \tag{3.13}$$

where the color indices a, b are in the adjoint representation of $SU(3)$.

Mass-terms are introduced to the SM Lagrangian through the Higgs doublet ϕ. The kinetic and potential terms of the Higgs can be written as

$$\mathcal{L}_H = (D_\mu \phi)^\dagger D^\mu \phi - V(\phi) \,, \tag{3.14}$$

where the potential is given by $V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2$. The Higgs undergoes spontaneous symmetry breaking (SSB), due to the non-zero vacuum-expectation value $\left\langle \phi^\dagger \phi \right\rangle = \frac{v^2}{2}$. In the unitary gauge [1], ϕ can be expanded around $v = \sqrt{\frac{\mu^2}{2\lambda}}$ as

$$\phi \to \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{3.15}$$

Inserting the expression back into $\mathcal{L}_H$ yields mass terms for electroweak gauge bosons. The mass eigenstates of the gauge bosons are defined through a rotation with the weak mixing angle θ_W:

$$\begin{pmatrix} \gamma \\ Z \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix}, \tag{3.16}$$

[1] The unitary gauge is non-renormalizable at higher orders, but we will only consider perturbative corrections in the QCD coupling constant.

and the charged states of W^1 and W^2 can be projected out with the linear combinations given by

$$W^{\pm} = \frac{1}{\sqrt{2}} \left(W^1 \mp i W^2 \right) , \qquad Q_0 \, W^{\pm} = \pm W^{\pm}. \tag{3.17}$$

The masses of the bosons after SSB are given in terms of the vacuum expectation value of the Higgs boson and the two coupling constants of the electroweak sector. They are given by

$$m_W = \frac{gv}{2} , \qquad m_Z = \frac{\sqrt{g^2 + g'^2}\, v}{2} , \qquad m_\gamma = 0 . \tag{3.18}$$

Additionally, the mixing angle can be expressed through the masses with $\cos\theta_W = \frac{m_W}{m_Z}$. The fermion masses are introduced through the same mechanism. The Yukawa Lagrangian, which contains interaction terms between the Higgs doublet and the fermions, can be written as

$$\mathcal{L}_{\text{Yukawa}} = - \sum_{i,j=1}^{3} \left(\bar{Q}_i \, Y_u^{ij} \, \tilde{\phi} \, u_j \right) - \sum_{i,j=1}^{3} \left(\bar{Q}_i \, Y_d^{ij} \, \phi \, d_j \right) - \sum_{i,j=1}^{3} \left(\bar{L}_i \, Y_e^{ij} \, \phi \, e_j \right) + \text{h.c.} , \tag{3.19}$$

where $\tilde{\phi} = i\sigma^2 \phi^*$ and Y_u^{ij}, Y_d^{ij} and Y_e^{ij} are complex 3 by 3 matrices that contain the coupling strength between fermion generations i and j. After SSB the Yukawa Lagrangian develops a mass matrix of the type $M_{u_i} = \frac{v}{\sqrt{2}} Y_{u_i}$, which is non-diagonal in flavor basis of the fermions. Diagonalising the states with unitary rotation yields a diagonal mass matrix and introduces the Cabibbo-Kobayashi-Maskawa matrix (CKM) in the quark sector. When massive neutrinos are considered one further introduces the Pontecorvo–Maki–Nakagawa–Sakata matrix (PMNS matrix) in the lepton sector, which describes neutrino mixing effects. The left-handed down-type quarks are replaced according to

$$\begin{pmatrix} d'_L \\ s'_L \\ b'_L \end{pmatrix} = V_{\text{CKM}} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} , \tag{3.20}$$

where the magnitude of the CKM matrix elements are measured experimentally to be [23]

$$V_{\text{CKM}} = \begin{pmatrix} 0.97435 \pm 0.00016 & 0.22501 \pm 0.00068 & 0.003732^{+0.000090}_{-0.000085} \\ 0.22487 \pm 0.00068 & 0.97349 \pm 0.00016 & 0.04183^{+0.00079}_{-0.00069} \\ 0.00858^{+0.00019}_{-0.00017} & 0.04111^{+0.00077}_{-0.00068} & 0.999118^{+0.000029}_{-0.000034} \end{pmatrix} , \tag{3.21}$$

which is close to unity for many practical purposes, including this thesis. We note in particular, that the top quark has only very weak mixing with the first two generations. Finally, the full SM Lagrangian at NLO in QCD is given by

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EW}} + \mathcal{L}_{GF} + \mathcal{L}_{\text{Ghost}} + \mathcal{L}_H + \mathcal{L}_{\text{Yukawa}} . \tag{3.22}$$

3.2 Effective Field Theories

Effective field theory (EFT) is the tool of choice to describe physics at a low energy scale while incorporating the effects of a theory realized at some much higher energy scale. Key to the concept is the observation that the high energy (small distance) physics should only provide a small modification to processes at much lower energies (large distances). This hierarchy allows for the consistent inclusion of the high energy effects as part of a systematic expansion. The expansion parameter is given by the ratio of the characteristic energy scales of the two theories. Consequently, the validity of the resulting theory is restricted to the regime, where the energies are much smaller than the high scale. The generating functional of the effective theory can be computed by integrating out the heavy degrees of freedom in the full theory. Schematically, for a heavy field ϕ_H and a light field ϕ_L the generating functional can be written as

$$\begin{aligned} Z[J_L, J_H] &= \int \mathcal{D}\phi_L \, \mathcal{D}\phi_H \, \exp\left(i \int d^4x \left(\mathcal{L}[\phi_L, \phi_H] + J_H \, \phi_H + J_L \, \phi_L\right)\right) \\ &\xrightarrow[\text{integrate out } \phi_H]{J_H = 0} \int \mathcal{D}\phi_L \exp\left(i \int d^4x \left(\mathcal{L}_{\text{EFT}}[\phi_L] + J_L \phi_L\right)\right) \equiv Z[J_L] \,, \end{aligned} \tag{3.23}$$

where the full theory Lagrangian is given by $\mathcal{L}$ and the effective Lagrangian by $\mathcal{L}_{\text{EFT}}$. The effective Lagrangian is expressed through the parameters of the full Lagrangian by matching their results at a chosen matching scale. The two theories described by $\mathcal{L}$ and $\mathcal{L}_{\text{EFT}}$ are formally independent. Indeed, the dynamics of such an effective theory should not be very closely related to those of the UV completion and their behavior in the high-energy regime will be different as a result. To ensure that the two theories describe the same physics in the low energy limit one has to perform a matching procedure, which fixes the free parameters of the effective theory in terms of those of the UV completion. Naively this can be achieved by computing S matrix elements in both theories, taking the infrared limit and solving for the parameters of the EFT. Schematically, one can write

$$\langle p_1, ..., p_n | \, S_{\text{UV}} \, | k_1, ..., k_m \rangle \overset{E/\Lambda \ll 1}{=} \langle p_1, ... p_n | \, S_{\text{EFT}} \, | k_1, ..., k_m \rangle \,, \tag{3.24}$$

where we have denoted the IR limit as $E/\Lambda \to 0$. An EFT of n free constants consequently requires the computation of n linearly independent scattering matrices to fully constrain the theory. Conveniently, this approach can be executed systematically order by order in perturbation theory. Beyond tree-level, UV singularities occur in both theories, which have to regularized separately with suitable renormalization counterterms. After UV renormalization this type of divergence cancels on each side of Eq. (3.24). At this point IR singularities remain on both sides of the matching equation [153]. However, by construction, the effective field theory exactly reproduces IR limit of the UV completion. As a result, all IR divergences must also cancel between both scattering-matrix expressions. Thus, for practical calculations one only needs to consider the finite part to extract the matching conditions [154]. At loop level, the full theory will develop logarithms of the kind $\log(\mu^2/\Lambda^2)$, where μ is the renormalization scale [2]. It is clear that this scale should be chosen close to Λ to avoid large logarithms, which could potentially spoil the perturbative convergence

[2] We assume dimensional regularization here, which is particularly convenient in EFT calculation.

of the calculation. The scale μ is called matching scale. The logarithms can not be reproduced by the dynamical fields in the EFT, because by definition no heavy degree of freedom is included. They are instead incorporated as finite higher order corrections to the fields and parameters in the effective field theory Lagrangian. Their implementation can be performed with a multiplicative approach similar to typical renormalization approaches.

The evaluation of scattering amplitudes at the loop level can be challenging in many respects. Naive application of Eq. (3.24) requires one to retain all masses in the higher-order calculation of the UV completion. One especially wants to avoid this when using the EFT as a means for approximation. Instead, it would be preferable to perform the expansion already at the integrand level. Doing so naively is not possible, because the loop momentum is integrated and thus assumes values, which are large compared to the masses of the full theory. However, more sophisticated approaches can still enable the desired simplifications. The so-called methods of regions [155–157] is used for this purpose in matching calculations. One first identifies the different kinematic regions of the loop-integral in the full theory in terms of their scaling behaviour. One then expands the integrand separately in all these regions, performs the integration and sums the results. One can realize that the expansion of the propagators of the full theory in the low momentum regime $k \ll \Lambda$, generates the corresponding integral in the effective field theory. On the other hand, the expansion in the high energy limit $k \gg m$, where m is the mass scale of the dynamical fields in the EFT, generates the correction that must be absorbed in the matching procedure. We can understand this with the following considerations. The EFT, by construction, replicates the IR behaviour of the underlying UV theory. Non-analytic terms in m are common to both theories. Since the matching at loop-level depends only on the difference between integrals in the two calculations, we can thus safely remove them before performing the actual integrals. This can be achieved by expanding the propagators in the EFT and UV theory with respect to the small mass m. However, the EFT never replicates any non-analytic terms of the UV theory, because the propagators have already been expanded. After expanding also the IR scales one is left only with scaleless integrals, which vanish in dimensional regularization [3]. We conclude that the difference of the integrals in the two theories can be computed by expanding the integrand in the full theory in $k \gg m$, which corresponds to the aforementioned high-energy region [158]. The region, where $m \ll k \ll \Lambda$, which one might think is missing in this discussion, is exactly the integral in the EFT expanded in m, which vanished in dimensional regularization. Consequently, one massively simplifies the problem by only expanding the UV theory integrals in $k \gg m$ and dropping the loop-correction to the EFT in the matching. The remaining integral no longer contains widely separated mass scales. Only the UV mass scale Λ is still present and generates the required non-analytical logarithmic terms.

While the matching procedure using the explicit computation of S-matrix elements is conceptually simple, one must in practice compute a very large number of matrix elements. Alternative matching approaches often aim to reduce the computational load, while also making the gauge invariance of the effective theory manifest at intermediate steps of the matching calculation. This can be achieved with the use

[3] This conclusion indeed depends on the choice of the regulator. The use of dimensionful regulators introduces additional massive scales and the scaleless integrals no longer occur.

of functional methods in the path integral formalism. One such approach is based on the *covariant derivative expansion* [159–161], which performs the matching at the level of the action instead of computing the scattering matrix elements. Within this technique, intermediate steps are always expressed in terms of gauge covariant expressions such as the covariant derivative. The action of the effective theory S_{eff} is evaluated to one-loop level by expanding the action of the UV completion S around the classical configuration given by

$$\left.\frac{\delta S[\phi,\Phi]}{\delta\Phi}\right|_{\Phi=\Phi_c} = 0\,, \tag{3.25}$$

where we follow the example of one light and heavy scalar field ϕ and Φ as in Ref. [154, 161]. Expanding the action around this minimum value one obtains

$$S_{eff}[\phi] \simeq S[\Phi_c] + ic_s \mathrm{Tr}\log\left(-\left.\frac{\delta^2 S}{\delta\Phi^2}\right|_{\Phi_c}\right)\,, \tag{3.26}$$

where we have performed the emerging Gaussian path integral analytically. The constant c_s generally depends on the spin representation of the respective heavy field. The first term is just the tree-level effective action, which can be obtained by solving the equation of motion for the heavy field and inserting it back into the action. To recover a polynomial in the light fields one then expands in terms of $1/M$, where M is the mass of the heavy field. The order at which one truncates this expansion corresponds to the mass-dimension of the generated operators.

To compute the one-loop contribution to the effective action one must evaluate the trace in Eq. (3.26). The form of the argument is given by

$$\left.-\frac{\delta^2 S}{\delta\Phi^2}\right|_{\Phi_c} = D^2 + M^2 + U(x)\,, \tag{3.27}$$

where $U(x)$ is a polynomial of light fields interacting with the heavy field Φ. The trace can be performed by inserting a complete basis of position and momentum states. After applying the Baker-Campbell-Hausdorff formula and one eventually arrives at

$$\begin{aligned}\Delta S_{eff} &= ic_s \frac{1}{(2\pi)^4}\int d^4x \int d^4p\, \mathrm{tr}\log\left(-\left(p_\mu + \tilde{G}_{\nu\mu}\frac{\partial}{\partial p_\nu}\right)^2 + M^2 + \tilde{U}\right)\,,\\ &\equiv \int d^4x\, \Delta\mathcal{L}_{eff}\end{aligned} \tag{3.28}$$

where $\tilde{G}$ and $\tilde{U}$ are defined as a tower of commutators by

$$\begin{aligned}\tilde{G}_{\nu\mu} &= \sum_{n=0}^{\infty}\frac{n+1}{(n+2)!}\left[-iD_{\alpha_1},[-iD_{\alpha_2},[...,[-iD_{\alpha_n},-iG_{\mu\nu}]]]\right]\prod_{i=1}^{n}\frac{\partial}{\partial p_{\alpha_i}}\,,\\ \tilde{U} &= \sum_{n=0}^{\infty}\frac{1}{n!}\left[-iD_{\alpha_1},[-iD_{\alpha_2},[...,[-iD_{\alpha_n},U]]]\right]\prod_{i=1}^{n}\frac{\partial}{\partial p_{\alpha_i}}\,,\end{aligned} \tag{3.29}$$

and the trace runs over the internal indices of the fields. The constant c_s depends on the spin representation of the fields. Each additional commutator in Eq. (3.29)

increases the dimension of the respective operator in the final expression. Instead of computing the logarithm, one inserts a derivative and integration in M^2 into Eq. (3.28), which yields

$$\Delta\mathcal{L}_{eff} = ic_s \frac{1}{(2\pi)^4} \int d^4p \int dM^2 \,\mathrm{tr} \frac{1}{-\Delta\left[1 + \frac{1}{\Delta}\left(\left\{p^\mu, \tilde{G}_{\nu\mu}\frac{\partial}{\partial p_\nu}\right\} + \tilde{G}_{\sigma\mu}\tilde{G}^\sigma_\nu \frac{\partial}{\partial p_\mu}\frac{\partial}{\partial p_\nu} - \tilde{U}\right)\right]}\,, \tag{3.30}$$

where $\Delta = \frac{1}{p^2 - m^2}$ is the free massive propagator. The final expression is obtained by expanding in terms of Δ to the desired order in the EFT expansion. Generalizations of this technique have since been developed [162–166].

Effective field theories constructed in this way have many important applications within particle physics. To just sketch a few:

- Theories with widely separated scales develop logarithms of the form $\log \frac{m}{M}$ in higher order calculations, where m is the mass of the light degree of freedom and M the heavy mass. Such logarithms can spoil perturbative convergence of the prediction, when $\alpha \log(\frac{m}{M}) \gtrsim 1$, where α is the expansion parameter of the perturbative series. The situation can be resolved using EFT techniques. First one integrates out the heavy degree of freedom, keeping only the dynamics of the light fields. The generated Wilson coefficients can be related at different energy scales through renormalization-group equations (RGE). The aforementioned large logarithms can be resummed through the solution of the RGE, which provides a all-order solution to their contribution.

- One can also employ the top-down approach as means of approximation. Given a field with a high mass M, one can employ the limit of $M \to \infty$ and perform a calculation instead in the simpler effective theory. EFTs provide the correct theoretical description of such an approximation. Examples of this approach include calculations in the heavy top-quark limit. At LHC the production of a single Higgs is loop-induced through a top-quark loop, which requires the calculation of loop-diagrams already at LO. The heavy-top limit replaces this loop by an effective operator, effectively reducing the loop-order by one.

While this *top-down* construction of EFTs is useful, one can also perform a *bottom-up* formulation. Because the EFT is itself a valid theory, which is not bound to a specific UV completion, one can directly use it. Instead of requiring knowledge of the full underlying theory, one then simply provides the most general form of Lagrangian in terms of the light degrees of freedom. The parameters of the Lagrangian can then be measured in experiments and the results interpreted by matching the EFT to the specific high-energy theories. A historical example is the Fermi-theory that incorporates the weak interaction in the form of four-fermion contact terms. From a modern perspective the Lagrangian of the weak Fermi theory can be written as

$$\mathcal{L}_{Fermi} = \frac{-g^2}{8m_W^2}\left(J^C_\mu\right)^\dagger J^{C,\mu} - \frac{g^2}{8m_W^2}\left(J^N_\mu\right)^\dagger J^{N,\mu}\,, \tag{3.31}$$

where the respective charged and neutral currents are given by

$$\begin{aligned} J^C_\mu &= \sum_f \frac{2I_3^f + 1}{2}\,\bar{f}\,\gamma_\mu\,(1-\gamma_5)\,f'\,, \\ J^N_\mu &= \sum_f \bar{f}\,\gamma_\mu\,(v_f + a_f\,\gamma_5)\,f\,, \end{aligned} \tag{3.32}$$

with the sum running over the fermion flavors, f' being the doublet partner of f [167] and v_f and a_f being vector and axial couplings to the Z-boson of the respective fermion f. One can observe that the operators are indeed suppressed with a massive scale $\frac{1}{m_W^2}$. The decay of a muon, $\mu^- \to e^- \nu_e \bar{\nu_\mu}$, is directly incorporated by this Lagrangian. Historically, the charged current term was used to describe the β-decay, $p \to n + e + \nu_e$, where (p, n) plays the role of one of the isospin doublets. The constants of such a term can be collected into the well-known Fermi constant G_F and experimentally determined. Indeed, in the original formulation the spinor structure did not restrict the theory to left-handed chiralities. However, this structure could be extracted from experimental observations. From the perspective of the *bottom-up* approach it is also unclear if both the β-decay and the muon-decay couple with the same coupling constant G_F. In retrospect, we know that both interactions are driven by a W-boson exchange in the underlying SM Lagrangian and thus expect the same interaction strength. Finding such a correspondence is itself a strong hint at a underlying symmetry in the full UV theory. Knowing the SM, we can compute the value of G_F by requiring that the matrix element computed within the SM agrees with the effective theory in the $m_W \to \infty$ limit, which implies vanishing momentum inside the W propagator at LO. In this thesis, we are particularly interested in this *bottom-up* approach. The general expression of any effective field theory Lagrangian is given by

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}^{ren.} + \sum_{n=5}^{N} \sum_{i} \frac{1}{\Lambda^{n-4}} C_i^{(n)} \mathcal{O}_i^{(n)} + \mathcal{O}\left(\frac{1}{\Lambda^{N+1}}\right) , \tag{3.33}$$

where $\mathcal{L}^{ren.}$ is a renormalizable Lagrangian that is build purely from at most four-dimensional, and thus renormalizable, operators. The scale Λ plays the role of the high energy scale used in the expansion. The sum over i runs over all operators $\mathcal{O}_i^{(n)}$ that can be build from the light degree of freedom of the renormalizable Lagrangian at a given power n in the expansion. Each operator is accompanied by a so-called Wilson coefficient C_i, which characterizes the interaction strength of the respective operator. The additional non-renormalizable operators are suppressed by the characteristic mass scale Λ, which dampens their impact on observables with increasing mass dimension. One can thus truncate the EFT expansion at a finite order N and obtain a predictive theory. In general, operators $\mathcal{O}_i$ can be related through various means and one should reduce the sum to only a non-redundant set to avoid flat directions in the coefficient space. Such a set is called a operator basis and is key to working with EFTs in the *bottom-up* approach. While the higher dimensional operators spoil the closed form of the renormalization in the traditional sense, one can still fully absorb all UV singularities at any finite order in the EFT expansion, which keeps the theory predictive in practical calculations. The *bottom-up* approach exploits that Eq. (3.33) requires no knowledge of the actual underlying Lagrangian at the high energy scale. Rather than Eq. (3.23) one then uses the Eq. (3.33) as the defining equation of the theory. To that end, one systematically constructs all operators $\mathcal{O}_i$ consistent with a chosen set of symmetries. The determination of the Wilson coefficients can be treated as parameters to be measured in experimental analyses.

3.3 Renormalization and operator mixing

The renormalization of the EFT can be constructed independently of any UV completion and is performed order by order in $1/\Lambda$. In that sense, the EFT is a fully consistent, predictive theory that can be used on its own. We focus in particular on the renormalization at the one-loop level here. We write the Lagrangian in terms of bare fields and parameters as

$$\mathcal{L}_{EFT} = \mathcal{L}^{ren.} + \sum_{n=5}^{N} \sum_{i_n} \frac{1}{\Lambda^{n-4}} C_{i_n}^{(n),B} \mathcal{O}_{i_n}^{(n)}(\{\phi_k^B\}, \{\lambda_k^B\}) + \mathcal{O}\left(\frac{1}{\Lambda^{N+1}}\right), \tag{3.34}$$

where the sets of bare fields and coefficients is denoted by $\{\phi_k^B\}$ and $\{\lambda_k^B\}$, respectively. The bare Wilson coefficients of the n-dimensional operators are given by $C_i^{(n),B}$. The renormalization of the bare fields and coefficients is performed using multiplicative renormalization, that is by introducing renormalized fields using $\phi_k^B = Z_{\phi_k}\phi_k$ and renormalized couplings using $\lambda_k^B = Z_{\lambda_k}\lambda_k$, as is standard in QFT. The higher dimensional operators generate additional singularities in loop calculations, which have to be absorbed into the definition of the Wilson coefficients. To that end, we define renormalized Wilson coefficients via

$$C_i^B = Z_{ij} C_j\,, \quad Z_{ij} = \delta_{ij} + \delta Z_{ij} \tag{3.35}$$

where Z_{ij} are the renormalization constants. Inserting effective operators into one-loop amplitudes generally does not yield structures proportional to the respective operators. This effect is called operator mixing. An example is given in Figure 3.1.

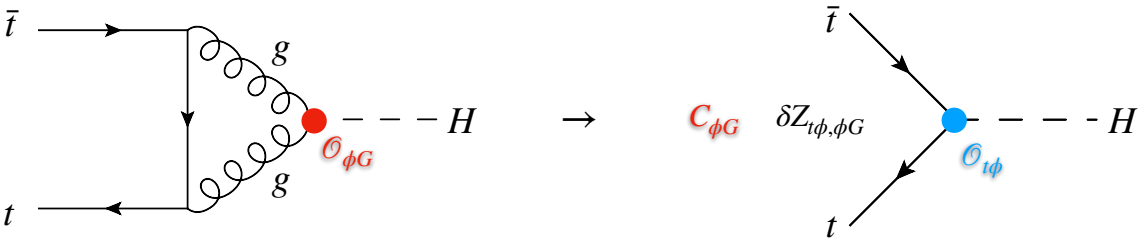


Figure 3.1: *Representative example of operator mixing. The operator $\mathcal{O}_{\phi G}$ mixes into $\mathcal{O}_{t\phi}$ and requires a counter-term proportional to the latter to renormalize. Taken from Ref. [3].*

The insertion of

$$\tilde{\mathcal{O}}_{\phi G} = \left(\phi^\dagger \phi\right) G^a_{\mu\nu} G^{a\mu\nu} \tag{3.36}$$

into the one-loop amplitude can be renormalized by a counterterm that is proportional to

$$\tilde{\mathcal{O}}_{t\phi} = \left(\phi^\dagger \phi\right) \bar{Q}\, t\, \tilde{\phi} + h.c.\,. \tag{3.37}$$

This can be facilitated with the non-diagonal entries of δZ_{ij}. A complete basis of operators can always accommodate any structure that might appear, as it would not be complete otherwise [4]. Practical calculations typically will only use a subset of operators and must ensure that all relevant contributions for the cancellation of

[4] Starting from two-loop calculations, the basis must generally be complete in d dimensions, which requires the addition of *evanescent operators*.

UV divergences are included. Furthermore, one must make sure to be consistent in the power counting of the high energy mass scale Λ. For example, renormalizing an EFT at dimension-8 requires not only the single insertions of dimension-8 operators, but also double insertions of dimension-6 operators. Similarly, if we were to consider multiple insertions of dimension-6 operators into the loop-amplitude then dimension-8 operators might also be required to obtain finite results. The exact values of δZ_{ij} must be extracted based on the pole-structure of the effective field theory in question. Special care needs to be taken when shifting operators by four-dimensional pieces. For example, the operators $\tilde{\mathcal{O}}_{\phi G}$ introduces a four-dimensional piece proportional to $G^a_{\mu\nu}G^{a\mu\nu}$, which breaks the canonical form of the Lagrangian. It is possible to shift the fields to restore the correct normalization of the kinetic term. However, it is easier to directly remove these modifications from the operator definition. For example, introducing

$$\mathcal{O}_{\phi G} = \left(\phi^\dagger\phi - \frac{v^2}{2}\right) G^a_{\mu\nu}\, G^{a\mu\nu}\,, \tag{3.38}$$

gets rid of any four-dimensional component in $\tilde{\mathcal{O}}_{\phi G}$. A subtlety arises when performing modifications of this kind. Operator mixing can require the introduction of these four-dimensional pieces in the renormalization of other operators. Removing these terms in the operator definition thus spoils the pole-cancellation. However, since these four-dimensional pieces are already present in the renormalizable part of the Lagrangian, we can recover the respective renormalization counterterms by adding a higher dimensional component to the renormalization of the fields ϕ_k and parameters λ_k. Schematically, we write the respective renormalization constants as

$$Z_i = 1 + \delta Z_i^{(4)} + \sum_{n=5}^{N} \delta Z_i^{(n)}/\Lambda^{n-4}\,, \tag{3.39}$$

where $\delta Z_i^{(n)}$ is a function of the Wilson coefficients $C_i^{(d)}$, and for each individual term the d of the coefficients sum to n. Alternatively, one can instead compute the renormalization terms using unshifted operators and then carefully shift the terms proportional to the desired Wilson coefficients afterwards. In dimensional regularization the logarithms generated in the one-loop calculation relate directly to the pole-structure. Consequently, the running of the renormalized Wilson coefficients relates to the pole-structure of the theory as well. The renormalization group equations (RGE) of the Wilson coefficients is given by

$$\frac{dC_i}{d\log\mu} = \gamma_{ij}C_j\,, \tag{3.40}$$

where we have introduced the anomalous dimension matrix γ_{ij}. Working in $d = 4-2\varepsilon$ dimensions and using the $\overline{\text{MS}}$ scheme we can relate the anomalous dimension matrix to the renormalization coefficients via

$$\delta Z_{ij} = \frac{1}{\varepsilon}\Delta(\mu_R, \mu_{EFT})\frac{1}{2}\gamma_{ij}\,, \tag{3.41}$$

where we have introduced

$$\Delta(\mu_R, \mu_{EFT}) = \Gamma(1+\varepsilon)\left(\frac{4\pi\mu_R^2}{\mu_{EFT}^2}\right)^{\varepsilon}\,, \tag{3.42}$$

to separate the scale μ_{EFT} at which the operators are defined from the renormalization scale μ_R. With this definitions the running between μ_R and μ_{EFT} is baked into the counterterms, which means that the Wilson coefficients are always defined at μ_{EFT}. This scheme is particularly useful when using dynamical scale definitions for μ_R as it allows us to factor out a constant value of $C_i(\mu_{EFT})$ from the amplitudes. The solution to the renormalization group equation of Eq. (3.40) can formally be written as

$$C_i(\mu) = \Gamma_{ij}(\mu, \mu_0)\, C_j(\mu_0)\,, \tag{3.43}$$

which relates the numerical value of the Wilson coefficients at two scales μ and μ_0. Practical solutions can be derived using various approximations or numerical methods. Focusing on the QCD running, we can expand the matrix γ_{ij} in powers of the strong coupling constant α_s as

$$\gamma_{ij} = \sum_{k=1}^{\infty} \left(\frac{\alpha_s}{4\pi}\right)^k \gamma_{ij}^{\mathrm{QCD},k}\,, \tag{3.44}$$

and truncate the series after the first term. The one-loop accurate result for Γ_{ij} is then given by

$$\begin{aligned} \Gamma^{\mathrm{QCD},1}(\mu, \mu_0) &= \exp\left(\int_{\mu_0}^{\mu} \frac{\alpha_s(\mu')}{4\pi\mu'} d\mu' \gamma^{\mathrm{QCD},1}\right) \\ &= \exp\left(\frac{1}{2\beta_0^{(n_F)}} \log\left(\frac{\alpha_s(\mu_0)}{\alpha_s(\mu)}\right) \gamma^{\mathrm{QCD},1}\right)\,, \end{aligned} \tag{3.45}$$

where the one-loop beta function coefficient $\beta_0^{(n_F)} = 11 - 2/3\, n_F$ and the number of flavors included in the α_s running is n_F. For the LHC application in this thesis, we will use the five-flavor scheme, i.e. $n_F = 5$.

Beyond the usual difficulties of handling the d-dimensional extension of four-dimensional objects like γ_5 one additional challenge arises in EFT calculations when using dimensional regularization. The typical operator basis are constructed in four dimensions and employ four-dimensional relations like Fierz-identities and γ-matrix relations to reduce their basis to a non-redundant set. However, the corresponding equalities often don't hold in non-integer dimensions. Important to SMEFT in this context are the four-fermion operators, whose spinor-product bilinears can be reduced using Fierz identities. Indeed, the operator basis in d dimensions is infinite in size, although only a finite number of terms contribute to a truncated perturbative calculation [168–170]. The additional operators required to construct a full basis are called *evanescent operators*. They vanish in the $\varepsilon \to 0$ limit and can be neglected in tree-level calculations. However, within higher order calculations they can contribute finite pieces, when combined with the UV pole of a loop integral. Indeed, a shift in the ε-dimensional basis results in a shift of the renormalized amplitudes of the physical operators. Their exact form depends on the way the d-dimensional structures are projected onto the four-dimensional operator basis. Starting at the two loop-level, the renormalization of a given four-dimensional operator can require counterterms, which are themselves part of the evanescent operator basis. This is clear as the operator basis would otherwise not be complete in d-dimensions. At the one-loop level this can not occur, because the coefficients of the UV poles must

be four-dimensional objects. We work at one-loop accuracy in this thesis and thus do not need to include them to obtain finite renormalized predictions. However, in general, an evanescent operator basis might still be required to allow for consistent matching calculations. As already mentioned, the choice of a basis of these operators is arbitrary and constitutes an additional scheme dependence of the calculation. Indeed, any evanescent operator E can be shifted by multiplying one of the four-dimensional operators multiplied by ε [169]. The dependence on the selected basis drops out when matched to a UV completion as long as the same convention is used [105]. Because the direct contributions of the evanescent operators always enter together with poles in one-loop calculations, one can always compensate their values with local counterterms [170]. Such a renormalization scheme is called *evanescent-free renormalization schemes*. It has been shown that one can impose renormalization conditions such that the anomalous dimension matrix contains a zero block for the mixing from evanescent operators to the physical basis [169, 171]. While a finite renormalization contribution persists, one can decouple the Wilson-coefficents of the evanescent operators from the RGE running of the physical Wilson coefficients. Given a specific choice of evanescent operators, one can always perform a matching calculation to a UV completion so that the scheme dependence drops out in physical predictions once all effects are included [105].

3.4 The SMEFT Lagrangian

SMEFT is the EFT extension to the SM. One assumes that, at the electroweak scale associated with the Higgs vacuum expectation value $v \approx 246\,\text{GeV}$, the dynamical field content is identical to the SM. The heavy degrees of freedom are characterized by the mass scale $\Lambda \gg v$. However, all the generated operators still realize the SM gauge group $SU(3) \times SU(2) \times U(1)$. The EFT operators are defined purely in terms of the SM field content. The renormalizable part of the Lagrangian is given by the well-known SM Lagrangian. The resulting SMEFT Lagrangian is given by

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{1}{\Lambda} C_i^{(5)} \mathcal{O}_i^{(5)} + \sum_i \frac{1}{\Lambda^2} C_i^{(6)} \mathcal{O}_i^{(6)} + \mathcal{O}\left(\frac{1}{\Lambda^3}\right), \tag{3.46}$$

where the bare Wilson coefficients are denoted $C_i^{(n)}$. SMEFT implements the spontaneous symmetry breaking (SSB) in the Higgs sector $SU(2) \times U(1) \to U(1)$ exactly as in the SM. However, the presence of the additional operators can modify the structure of the Higgs potential, which then has to be be respected in the derivation of the vacuum expectation value. For example, at dimension-6 the operator $\mathcal{O}_\phi \equiv \left(\phi^\dagger\phi - \frac{v^2}{2}\right)^3$ modifies the overall potential to

$$V(\phi^\dagger\phi) = -\mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2 + \frac{C_\phi}{\Lambda^2}\left(\phi^\dagger\phi - \frac{v^2}{2}\right)^3. \tag{3.47}$$

This placed the new vacuum expectation value at

$$\left\langle \phi^\dagger\phi \right\rangle = \frac{v^2}{2}\left(1 + \frac{3C_\phi v^2}{4\lambda\Lambda^2}\right), \tag{3.48}$$

where v is the SM vacuum expectation value and the exact solution of the vacuum expectation value was expanded in terms of C_ϕ/Λ^2 to first order [154], which is

valid when $C_\phi \left\langle \phi^\dagger \phi \right\rangle / \Lambda^2 \ll 1$. The relations among other SM constants and input parameters are similarly modified by the additional operators.

The Wilson coefficients are treated as free parameters that can be constrained by experiments, but can in principle also be matched to from an explicit UV theory. Both approaches can be utilised in the search for new physics at the LHC. The former directly constrains potential operators that could be used to make predictions if they are found to be non-vanishing. The latter strategy is particularly useful to efficiently probe the viability of a proposed BSM model. Knowing the experimental constrains on the SMEFT Wilson coefficients allows one to directly test a potential BSM model through a matching calculation without requiring a dedicated experimental study for each proposed BSM candidates. Investing effort into the theoretical development of SMEFT thus is not only useful for the search of new physics in a model independent way, but also complements the search for explicitly formulated BSM models. In fact, even if such a model would be found to be realized in nature, one could still utilize SMEFT to efficiently compute predictions at the electroweak scale. Consequently, the pursuit of SMEFT is likely rewarded in a multitude of potential BSM scenarios [154].

While the formulation seems deceptively simple, the actual number of operators can increase very quickly with the mass-dimension and the number of fermion generations under consideration. Furthermore, special care needs to be taken to only include non-redundant operators. Operators combination proportional to equations of motion can be removed or equivalently eliminated through a shift of integration variable in the path integral formalism. Because total derivatives don't impact physical predictions one can relate operators that differ purely by one as well. Lastly, care must be taken to eliminate operators that can be related with Fierz-identities. The reduction to a fully non-redundant basis is non-trivial. Historically, multiple overcomplete bases were published. For example, the first complete non-redundant basis of SMEFT at dimension-6 reports 59 baryon-number conserving operators [172] for a single fermion generation, down from the 80 reported in [173]. In recent years much progress has been made in the automated derivation of these operator basis. Indeed, computer codes that automatically derive the basis up to arbitrary mass-dimension have been published in [174, 175], only limited by the computational resources required to perform the calculation. At dimension-5, neglecting flavor indices, the only additional operator is the well-known Weinberg operator [176], which can generate neutrino masses in the Lagrangian. The operator is given by

$$\mathcal{O}^{(5)} = \kappa_{ij} \left(\tilde{\phi}^\dagger L_i\right) C \left(\tilde{\phi}^\dagger L_j\right) + h.c.\,, \tag{3.49}$$

where κ_{ij} is a symmetric matrix connecting the different flavor generations and C is the charge conjugation matrix [176]. One can construct other contractions of the ϕ and L_i doublets. The above definition of the Weinberg operator corresponds to a product of two $SU(2)$ singlets $\tilde{\phi}^\dagger L_i$. First, we note that the other singlet contribution vanishes because $\tilde{\phi}^\dagger \phi = 0$ after SSB. One can construct products of two $SU(2)$ triplets instead. For example, one can write $\kappa_{ij}(\tilde{\phi}^\dagger \sigma^I L_i)\, C\, (\tilde{\phi}^\dagger \sigma^I L_j) + h.c.$ However, with the use of Fierz-identity

$$\sigma^I_{ab}\sigma^I_{cd} = 2\delta_{ad}\delta_{bc} - \delta_{ab}\delta_{cd}\,, \tag{3.50}$$

one can reduce this structure back to the singlet formulation [177]. This operator produces neutrino masses after SSB of the Higgs sector, which are proportional to

v^2/Λ. Consequently, for masses below the eV scale one must have very high scales Λ, which would make the corresponding new physics effects otherwise very difficult to access. This is known as the Seasaw-mechanism [178, 179].

The dimension-6 basis contains the aforementioned 59 operators for one fermion generation, where hermitian conjugated terms are not counted separately. Relaxing the restriction on baryon-number violation increases the count to 63 operators. Restoring three fermion generations increases the number to 2499 [113] and 3045 [175], respectively. This basis is the so-called Warsaw basis and has become the defacto standard in modern SMEFT calculations. Similarly, the basis of dimension-7 SMEFT was derived fully in [180], which eliminated two operators of the set reported in [181]. Notably, all operators at this order violate Lepton number conservation and several also violate Baryon number conservation. Modern automated algorithms were further used to derive the non-redundant basis of dimension 8 [182] and 9 [183]. Lastly, results up to dimension 12 were reported in [184]. The large number of free parameters makes working with the full SMEFT Lagrangian difficult in practice. SMEFT calculations thus often limit themselves to specific sectors of the theory or impose additional symmetries explicitly to reduce the computational complexity. The accidental lepton number and baryon number symmetries of the SM are often imposed explicitly on the effective operators. Similarly, $\mathcal{CP}$ violating contributions are often limited to dedicated studies so that general calculations of LHC processes can forego their inclusion [154].

The previously discussed dimension-5 Lagrangian primarily captures potential BSM effects in the neutrino sector. In order to study interactions of the top quark and the Higgs boson one thus has to work with at least the dimension-6 operators. On the other hand, working with higher mass dimensions increases the complexity of the calculation even further. One can explicitly require baryon-number and lepton-number conservation, which is equivalent to removing the 5 and 7 dimensional operators. Indeed, even the dimension-6 basis is very large when considering three fermion generations. There are three types of relevant operators for top-quark processes at the LHC:

- Operators that directly generate or modify top-quark interactions,
- operators that modify the existing SM top-quark coupling strength by changing the relation of the input constants and SM parameters,
- and operators that change the interactions of the initial state partons inside the top-quark production processes at the LHC.

In the Warsaw-basis, the first group contains 28 operators, while the second and third categories contain 5 and 4 operators, respectively. In full generality, these assumptions produce 1179 parameters out of which 557 are complex phases [154]. To simplify the picture one can make use of an approximation. Imposing a

$$U(2)_Q \times U(2)_u \times U(2)_d \times [U(1)_L \times U(1)_e]^3$$

symmetry on the effective operators enforces flavor universality in the lepton sector and restricts the quark interactions to be symmetric in the first two quark genera-

tions. One can go one step further by instead requiring a

$$U(2)_Q \times U(2)_u \times U(3)_d \times [U(1)_L \times U(1)_e]^3$$

flavor symmetry, which is consistent with the five-flavor scheme used in many LHC calculations. The former restrictions already reduce the number of free parameters to 85. This is still a rather large number, but one can make it manageable by identifying subsets of operators that are particularly well constrained in specific processes. By combining measurements of several observables sensitive to different operators, one can obtain a comprehensive picture without determining the full basis simultaneously. For example, Ref. [185] reports that the leading $\mathcal{CP}$-violating effects can be isolated to just 10 operators [5]. The latter symmetry group is used in practice within the NLO QCD UFO model SMEFT@NLO, which also imposes $\mathcal{CP}$-conservation. At NLO QCD, the unitary gauge is employed in the Higgs sector in this model. More general gauge choices have also been studied e.g. in Refs. [107, 186–191].

3.5 The Warsaw basis in the top-quark sector

Bosonic			
$\mathcal{O}_i$	Definition	$\mathcal{O}_i$	Definition
$\mathcal{O}_G$	$g_s f^{abc} G^a_{\mu\nu} G^{b\nu\rho} G^{c\mu}_{\rho}$	$\mathcal{O}_W$	$\epsilon_{IJK} W^I_{\mu\nu} W^{J\nu\rho} W^{K\mu}_{\rho}$
$\mathcal{O}_{\phi G}$	$\left(\phi^\dagger\phi - \frac{v^2}{2}\right) G^{a\mu\nu} G^a_{\mu\nu}$	$\mathcal{O}_{\phi W}$	$\left(\phi^\dagger\phi - \frac{v^2}{2}\right) W^{I\mu\nu} W^I_{\mu\nu}$
$\mathcal{O}_{\phi BB}$	$\left(\phi^\dagger\phi - \frac{v^2}{2}\right) B^{\mu\nu} B_{\mu\nu}$	$\mathcal{O}_{\phi WB}$	$\left(\phi^\dagger\sigma_I\phi\right) B^{\mu\nu} W^I_{\mu\nu}$
$\mathcal{O}_\phi$	$\left(\phi^\dagger\phi - \frac{v^2}{2}\right)^3$	$\mathcal{O}_{\phi d}$	$-\partial_\mu(\phi^\dagger\phi)\,\partial^\mu(\phi^\dagger\phi)$
$\mathcal{O}_{\phi D}$	$(\phi^\dagger D^\mu\phi)^\dagger(\phi^\dagger D_\mu\phi)$		

Table 3.1: *Bosonic operators relevant to the top-quark sector of SMEFT as implemented in the SMEFT@NLO UFO model.*

In this thesis, we adopt the conventions of the SMEFT@NLO model, which can be found on the corresponding webpage [6]. The full definitions of the four-fermion operators can be found in Ref. [192]. Further information was published alongside the model in Ref. [105], which also contains the definition of the evanescent operator basis as well as the d-dimensional Dirac algebra scheme. To recap, at NLO the

$$U(2)_Q \times U(2)_u \times U(3)_d \times [U(1)_L \times U(1)_e]^3$$

flavor symmetry is enforced on the effective operator basis. In addition, $\mathcal{CP}$ conservation is required. This selection of operators is constructed with particular focus

[5] This number increases to 17 when the bottom quark is treated as massive.

[6] `https://cp3.irmp.ucl.ac.be/projects/feynrules/wiki/SMEFTatNLO`

2 fermion (chiral flip)			
$\mathcal{O}_i$	Definition	$\mathcal{O}_i$	Definition
$\mathcal{O}_{t\phi}$	$\left(\phi^\dagger\phi - \frac{v^2}{2}\right)\bar{Q}\,t\,\tilde{\phi} + h.c.$	$\mathcal{O}_{tW}$	$i\left(\bar{Q}\sigma^{\mu\nu}\sigma^I t\right)\tilde{\phi}\,W^I_{\mu\nu} + h.c.$
$\mathcal{O}_{tG}$	$ig_s\left(\bar{Q}\sigma^{\mu\nu}t^a t\right)\tilde{\phi}\,G^a_{\mu\nu} + h.c.$	$\mathcal{O}_{tB}$	$i\left(\bar{Q}\sigma^{\mu\nu}t\right)\tilde{\phi}\,B_{\mu\nu} + h.c.$
2 fermion (current)			
$\mathcal{O}_i$	Definition	$\mathcal{O}_i$	Definition
$\mathcal{O}^{(1)}_{\phi L_1}$	$(\phi^\dagger \overleftrightarrow{D}_\mu \phi)(\bar{L}_1\gamma^\mu L_1)$	$\mathcal{O}^{(3)}_{\phi L_1}$	$(\phi^\dagger \overleftrightarrow{D}_\mu \sigma_I \phi)(\bar{L}_1\gamma^\mu \sigma^I L_1)$
$\mathcal{O}^{(1)}_{\phi L_2}$	$(\phi^\dagger \overleftrightarrow{D}_\mu \phi)(\bar{L}_2\gamma^\mu L_2)$	$\mathcal{O}^{(3)}_{\phi L_2}$	$(\phi^\dagger \overleftrightarrow{D}_\mu \sigma_I \phi)(\bar{L}_2\gamma^\mu \sigma^I L_2)$
$\mathcal{O}^{(1)}_{\phi L_3}$	$(\phi^\dagger \overleftrightarrow{D}_\mu \phi)(\bar{L}_3\gamma^\mu L_3)$	$\mathcal{O}^{(3)}_{\phi L_3}$	$(\phi^\dagger \overleftrightarrow{D}_\mu \sigma_I \phi)(\bar{L}_3\gamma^\mu \sigma^I L_3)$
$\mathcal{O}_{\phi e_1}$	$(\phi^\dagger \overleftrightarrow{D}_\mu \phi)(\bar{e_1}\gamma^\mu e_1)$	$\mathcal{O}_{\phi e_2}$	$(\phi^\dagger \overleftrightarrow{D}_\mu \phi)(\bar{e}_2\gamma^\mu e_2)$
$\mathcal{O}_{\phi e_3}$	$(\phi^\dagger \overleftrightarrow{D}_\mu \phi)(\bar{e_3}\gamma^\mu e_3)$		
$\mathcal{O}^{(1)}_{\phi q_i}$	$\sum_{i=1}^2(\phi^\dagger \overleftrightarrow{D}_\mu \phi)(\bar{q}_i\gamma^\mu q_i)$	$\mathcal{O}^{(1)}_{\phi Q}$	$(\phi^\dagger \overleftrightarrow{D}_\mu \phi)(\bar{Q}\gamma^\mu Q)$
$\mathcal{O}^{(3)}_{\phi q_i}$	$\sum_{i=1}^2(\phi^\dagger \overleftrightarrow{D}_\mu \sigma_I \phi)(\bar{q}_i\gamma^\mu \sigma^I q_i)$	$\mathcal{O}^{(3)}_{\phi Q}$	$(\phi^\dagger \overleftrightarrow{D}_\mu \sigma_I \phi)(\bar{Q}\gamma^\mu \sigma^I Q)$
$\mathcal{O}_{\phi u_i}$	$\sum_{i=1}^2(\phi^\dagger \overleftrightarrow{D}_\mu \phi)(\bar{u_i}\gamma^\mu u_i)$	$\mathcal{O}_{\phi d_i}$	$\sum_{i=1}^3(\phi^\dagger \overleftrightarrow{D}_\mu \phi)(\bar{d_i}\gamma^\mu d_i)$
$\mathcal{O}_{\phi u_i}$	$(\phi^\dagger \overleftrightarrow{D}_\mu \phi)(\bar{t}\gamma^\mu t)$		

Table 3.2: *2-fermion operators relevant to the SMEFT@NLO model. The operators are split into "chiral flip" and "current" type operators.*

on Higgs and top-quark studies at the LHC. The relevant operators under these assumptions are listed in the following. Purely bosonic operators are listed in Table 3.1, operators containing two fermions in Table 3.2 and the four-fermion operators are split into those containing two quarks and two leptons as listed in Table 3.3 as well as those with four-quarks in Table 3.4.

Notably, some operators coefficients are traded to isolate specific interaction vertices to individual Wilson coefficients or better match particular combinations that might be measured in common LHC processes. Indeed, a common framework given in Ref. [192] provides many such linear combinations of relevant operators. Adopting these reduces the amount of degenerate parameters in experimental analyses and can reduce the number of operator coefficients that have to be included in a given theoretical calculation. For example, the C_{tB} coefficient can be replaced using

$$C_{tZ} = -\sin\theta_W\, C_{tB} + \cos\theta_W\, C_{tW}\,, \tag{3.51}$$

which isolates the $t\bar{t}Z$ vertex to the C_{tZ} coefficient after SSB by cancelling the corresponding contribution proportional to C_{tW}. Another example are the $C^{(1)}_{\phi Q_i}$ and

2 lepton 2 quark			
$\mathcal{O}_i$	Definition	$\mathcal{O}_i$	Definition
$\mathcal{O}_{LQ}^{1(ijkl)}$	$(\bar{L}_i\gamma^\mu L_j)(\bar{Q}_k\gamma_\mu Q_l)$	$\mathcal{O}_{LQ}^{3(ijkl)}$	$(\bar{L}_i\gamma^\mu\sigma_I L_j)(\bar{Q}_k\gamma_\mu\sigma^I Q_l)$
$\mathcal{O}_{Lu}^{(ijkl)}$	$(\bar{L}_i\gamma^\mu L_j)(\bar{u}_k\gamma_\mu u_l)$	$\mathcal{O}_{eQ}^{(ijkl)}$	$(\bar{e}_i\gamma^\mu e_j)(\bar{Q}_k\gamma_\mu Q_l)$
$\mathcal{O}_{eu}^{(ijkl)}$	$(\bar{e}_i\gamma^\mu e_j)(\bar{u}_k\gamma_\mu u_l)$		
$\mathcal{O}_{LeQu}^{1(ijkl)}$	$(\bar{L}_i e_j)i\sigma^2(\bar{Q}_k u_l) + h.c.$	$\mathcal{O}_{LeQu}^{3(ijkl)}$	$(\bar{L}_i\sigma^{\mu\nu} e_j)i\sigma^2(\bar{Q}_k\sigma_{\mu\nu} u_l)$

Table 3.3: *Four-fermion operators with two leptons and two quarks relevant to the SMEFT@NLO model. The operators employed in the UFO model are given by specific subset of index combinations.*

4 quarks			
$\mathcal{O}_i$	Definition	$\mathcal{O}_i$	Definition
$\mathcal{O}_{QQ}^{1(ijkl)}$	$(\bar{Q}_i\gamma^\mu Q_j)(\bar{Q}_k\gamma_\mu Q_l)$	$\mathcal{O}_{QQ}^{3(ijkl)}$	$(\bar{Q}_i\gamma^\mu\sigma_I Q_j)(\bar{Q}_k\gamma_\mu\sigma^I Q_l)$
$\mathcal{O}_{Qu}^{1(ijkl)}$	$(\bar{Q}_i\gamma^\mu Q_j)(\bar{Q}_k\gamma_\mu Q_l)$	$\mathcal{O}_{Qu}^{8(ijkl)}$	$(\bar{Q}_i\gamma^\mu t^a Q_j)(\bar{u}_k\gamma_\mu t^a u_l)$
$\mathcal{O}_{ud}^{1(ijkl)}$	$(\bar{u}_i\gamma^\mu u_j)(\bar{d}_k\gamma_\mu d_l)$	$\mathcal{O}_{ud}^{8(ijkl)}$	$(\bar{u}_i\gamma^\mu t^a u_j)(\bar{d}_k\gamma_\mu t^a d_l)$
$\mathcal{O}_{uu}^{(ijkl)}$	$(\bar{u}_i\gamma^\mu u_j)(\bar{u}_k\gamma_\mu u_l)$		

Table 3.4: *Four-fermion operators with four quarks relevant to the SMEFT@NLO model. The operators employed in the UFO model are given by specific subset of flavor index combinations.*

$C_{\phi Q}^{(1)}$ parameters, which can be traded for the combinations

$$\left(C_{\phi Q_i}^{(1)} - C_{\phi Q_i}^{(3)}\right), \quad \text{and} \quad \left(C_{\phi Q}^{(1)} - C_{\phi Q}^{(3)}\right). \tag{3.52}$$

In the four-fermion sector only operators with at least one $t\bar{t}$ pair are retained. Consequently, at least two indices in $(ijkl)$ are always set to 3. Among the two-lepton two-quark operators the $C_{LQ}^{1,(ii33)}$ Wilson coefficient is replaced by the linear combination

$$C_{LQ}^{1,(ii33)} - C_{LQ}^{3,(ii33)}, \qquad i = 1, 2, 3, \tag{3.53}$$

where we have adopted the notation of the generation indices of the operators also for the Wilson coefficients. The $\mathcal{O}_{LeQu}$ operator appears only with indices (3333), but the remaining coefficients in this class carry indices $(ii33)$ with $i = 1, 2, 3$. To

summarize, the Wilson coefficients in this sector are chosen to be

$$\begin{aligned} &C_{LQ}^{1,(ii33)} - C_{LQ}^{3,(ii33)}\,, && C_{LQ}^{3,(ii33)}\,, \\ &C_{LeQu}^{1(3333)}\,, && C_{LeQu}^{3(3333)}\,, \\ &C_{Lu}^{(ii33)}\,, && C_{eu}^{(ii33)}\,, \\ &C_{eQ}^{(ii33)}\,, \end{aligned} \tag{3.54}$$

where i runs over all three lepton flavor generations. Finally the four-quark sector can be split into two categories. Those with all heavy quarks and those, which are mixed between heavy and light quarks. Those former are given by

$$\mathcal{O}_{Qu}^{8(3333)}\,, \quad \mathcal{O}_{uu}^{(3333)}\,, \quad \mathcal{O}_{Qu}^{1(3333)}\,, \quad \mathcal{O}_{QQ}^{1/3(3333)}\,, \tag{3.55}$$

where we used m/n as a shorthand for both index assignments. In the SMEFT@NLO conventions, the Wilson coefficients of the latter two operators are replaced with

$$\left(2\,C_{QQ}^{1(3333)} - \frac{2}{3}\,C_{QQ}^{3(3333)}\right) \quad \text{and} \quad 8\,C_{QQ}^{3(3333)}\,. \tag{3.56}$$

The mixed four-fermion operators contain operators with two heavy quark fields and two light ones. The symmetry group enforced in the model delivers all operators indexed with the first and second quark generation, in other words the following operators all appear with $i = 1, 2$. These include

$$\mathcal{O}_{ud}^{1/8(33ii)}\,, \quad \mathcal{O}_{Qu}^{1/8(ii33)}\,, \quad \mathcal{O}_{Qu}^{1/8(33ii)}\,, \quad \mathcal{O}_{ud}^{1/8(33ii)}\,. \tag{3.57}$$

Furthermore, the two operators $\mathcal{O}_{uu}^{(ii33/i33i)}$ are selected, where the coefficients are given by the linear combinations

$$C_{uu}^{(ii33)} + \frac{1}{3}\,C_{uu}^{(i33i)} \quad \text{and} \quad 2\,C_{uu}^{(i33i)}\,. \tag{3.58}$$

Lastly, the $\mathcal{O}_{QQ}^{1(ijkl)}$ and $\mathcal{O}_{QQ}^{3(ijkl)}$ operators are included with the two index assignments $(ii33)$ and $(i33i)$. To complete the definitions of the degrees of freedom one defines the combinations

$$\begin{aligned} &C_{QQ}^{1,(ii33)} + \frac{1}{6}\,C_{QQ}^{1,(i33i)} + \frac{1}{2}\,C_{QQ}^{3,(i33i)}\,, \\ &C_{QQ}^{1,(i33i)} + 3\,C_{QQ}^{3,(i33i)}\,, \\ &C_{QQ}^{3,(ii33)} + \frac{1}{6}\,C_{QQ}^{1,(i33i)} - \frac{1}{6}\,C_{QQ}^{3,(i33i)}\,, \\ &C_{QQ}^{1,(i33i)} - C_{QQ}^{3,(i33i)} \end{aligned} \tag{3.59}$$

as the free parameters of these operators.

The category of bosonic operators contains in particular $\mathcal{O}_G$ and $\mathcal{O}_W$, which modify the self-interactions of the gauge bosons and introduce additional interaction with up to six of them. $\mathcal{O}_G$ is normalized with an additional power of g_s, which unifies the power counting of the coupling constants between SM and SMEFT calculations. $\mathcal{O}_\phi$, $\mathcal{O}_{\phi d}$ and $\mathcal{O}_{\phi D}$ introduces modifications to the Higgs boson self-interactions. The latter in particular also generates Higgs - vector-boson interactions with up to six

legs. While these are not directly connected to top quarks, they do affect typical top-quark productions processes at the LHC or modify the relations among the SM parameters. However, in practical calculations it is often assumed that operators unrelated to the Higgs boson and the top quark can be constrained using other processes. This has been shown to be true at least in specific circumstances such as the case of $\mathcal{O}_G$, which has been studied in Ref. [193] within the multi-jet production at LHC. Particularly relevant for $t\bar{t}$-associated production and single-top processes is the $\mathcal{O}_{\phi G}$ operators, because it generates effective Higgs-gluon interactions, which contribute to the $pp \to t\bar{t}H$ process.

The two-fermion operators can be split into two categories based on their chirality structure. The "chiral flip" operators relate left- and right-handed quarks, while the "current" type retains the SM spinor structure. The model assumptions do indeed include a number of lepton-Higgs interactions in the latter category, which are not probed by top-quark physics. On the other hand, $\mathcal{O}_{\phi Q}^{(1/3)}$ can modify the top decay process $t \to Wb$, while the corresponding light-flavored operators $\mathcal{O}_{\phi q_i}^{(1/3)}$ can partake in quark-initiated partonic channels. However, most LHC processes are dominated by gluon initiated channels. Therefore, only reduced sensitivity to these is expected. The "chiral flip" operators are all important in top-quark physics. $\mathcal{O}_{t\phi}$ modifies the top-Higgs interaction, while adding additional vertices with extra Higgs bosons attached to the vertex. The two top-vector boson operators $\mathcal{O}_{tW}$ and $\mathcal{O}_{tB}$ enter both in the production process of associated productions like $pp \to t\bar{t}W$ and $pp \to t\bar{t}Z$. The former also enters the top-quark decay, where it provides a unique spin structure compared to the SM and $\mathcal{O}_{\phi Q}^{(1/3)}$. This makes the study of this operator particularly interesting. The top chromomagnetic operator $\mathcal{O}_{tG}$ modifies many experimental analyses related to $t\bar{t}$ production and is thus an important term to study. It modifies the top-gluon interaction, which is predominant in the production of $t\bar{t}$-associated processes, while also adding vertices with an additional attached Higgs leg. The latter of course are relevant to the $pp \to t\bar{t}H$ process, which we study in this thesis.

The four-fermion operators typically only enter through the quark initiated partonic processes. They are consequently suppressed due to the much smaller size of the quark PDFs compared to the gluons at LHC. However, the operators with at least two light flavor can enter already at LO in the perturbative expansion in α_s. On the other hand, the operators containing only the third quark generation require either four tops or at least two bottom quarks, which are restricted to the $pp \to t\bar{t}t\bar{t}$ and $pp \to t\bar{t}b\bar{b}$ processes. The latter case can also contribute through the suppressed bottom PDFs, although such contributions tend to be negligible in many top-quark processes. At NLO, this picture is slightly relaxed as the additional heavy lines can be contained within a loop at the cost of an additional suppression factor α_s.

Chapter 4

HELAC-NLO

The HELAC-NLO framework is a software package that enables the computation of SM cross sections for arbitrary processes at one-loop accuracy. The package consists of two main constituents, namely HELAC-1LOOP [127, 194] and HELAC-DIPOLES [195, 196]. The two codes are responsible for the computation of finite virtual and real corrections, respectively. Once a finite amplitude has been obtained one can integrate these separately using Monte Carlo integration. We employ the phase-space generator KALEU [197] for this purpose. The optimization of the grids is performed using the PARNI [198] package. We compute the necessary Born-level helicity amplitudes using the Dyson-Schwinger recursive algorithm as implemented in HELAC [199]. Finite real-emission contributions are computed in HELAC-DIPOLES using the subtraction schemes described in Section 4.3.1. The estimate of theoretical uncertainties typically requires the evaluation of seven scale setups and the evaluation of $\mathcal{O}(100)$ different PDF eigenvectors. The costly recomputation of the full integration can be avoided when different setups are reasonable similar to each other. In that case, one can store unweighted event files alongside additional information on each event to fully reconstruct the dependence on the scales and PDFs. These events can then be reweighted for the different setups in a fraction of the time. This functionality is implemented in the in-house software HEPLOT [200] . Finally, distributions are obtained by accumulating histogram entries of desired observables for each stored event. HELAC-1LOOP evaluates virtual corrections using the OPP method implemented in CUTTOOLS [201, 202]. The required scalar master integrals are provided in the ONELOOP [203] library. This approach can support arbitrary processes and thus lends itself to automation. The runtime for a single phase-space point exceeds the tree-level calculation by orders of magnitude. However, 1-loop virtual corrections are given by an interference of the Born and 1-loop matrix elements. Consequently, the shape of the virtual corrections is still broadly similar to those of the LO calculation. HELAC-1LOOP exploits this observation by unweighting the LO cross section first. These events are then reweighted to the corresponding virtual correction contribution, partially bypassing the expensive integration of the one-loop amplitudes. Finally, the events are then stored in a similar manner as in HELAC-DIPOLES to allow for further reweighting in HEPLOT. This chapter outlines the techniques used throughout the HELAC-NLO package. The extension and modification for the application to SMEFT processes is discussed separately in Chapter 5.

4.1 Born-level matrix element computations

4.1.1 Recursive computation of helicity amplitudes

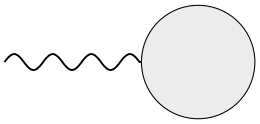
The standard approach to computing tree-level amplitudes based on Feynman diagrams is naturally suited to automated methods, yet its computational cost asymptotically grows factorially with the number of particles. High multiplicity final states thus become prohibitively expensive to compute [199, 204]. Alternatively, one can employ recursive algorithms based on the Dyson-Schwinger equations [199, 204–206]. These scale exponentially as a^n with $a \sim 3$ and are thus more efficient. In practical terms, the Dyson-Schwinger equations allow to recursively express n-point Green's functions in terms of $1, 2, ..., (n-1)$-point functions. It is convenient to categorize Green's functions in terms of the included external momenta. The subset of contained external momenta $I \in \{1, 2, ..., n\}$ can then be mapped to a unique integer

$$m(I) = \sum_{i \in I} 2^{i-1} \,, \tag{4.1}$$

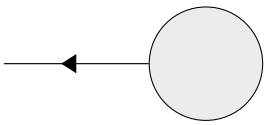
with the total contained momenta of the current being given by

$$P(I) = \sum_{i \in I} p_i \,. \tag{4.2}$$

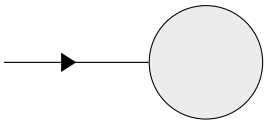
Implementations for general processes will need to track additional information on the computed (sub-)currents, such as flavor content and the number of perturbative couplings included. For simplicity, we illustrate the algorithm in the QED case following closely the original publication [199], where the momentum content is sufficient for categorization. We define vector currents with momentum P as

$A_\mu(P) =$  , (4.3)

the four-dimensional spinor corresponding to a fermion sub-amplitude as

$\psi(P) =$  , (4.4)

and finally the four-dimensional anti-spinor, which describes an anti-fermion, by

$\bar{\psi}(P) =$  . (4.5)

The Dyson-Schwinger equation for the vector boson can be written as

$$A^\mu(P) = \sum_{i=1}^{n} \delta_{P=p_i} A^\mu(p_i) + \sum_{P=P_1+P_2} (ie)\, \Pi^\mu_{\ \nu}\, \bar{\psi}(P_2)\gamma^\nu \psi(P_1)\epsilon(P_1, P_2) \,, \tag{4.6}$$

where e is the coupling constant, $\Pi^\mu_{\ \nu}$ is the propagator in R_ξ gauge, given by

$$\Pi^\mu_{\ \nu} = \frac{i}{P^2 - m^2} \left(-g^\mu_\nu + (1-\xi)\frac{P^\mu P_\nu}{P^2 - \xi m^2} \right) \,, \tag{4.7}$$

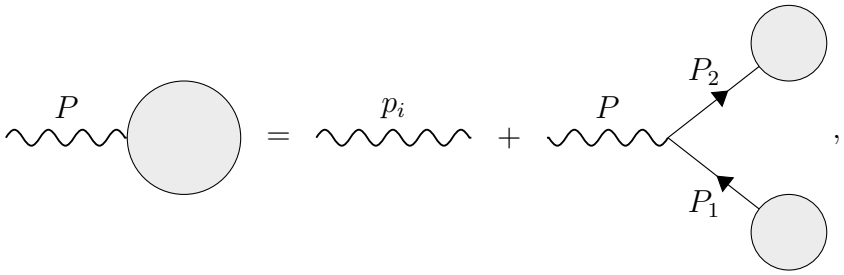
and $\epsilon(P_1, P_2)$ is an additional sign factor designed to keep track of fermion exchanges. The factor can be computed using the binary representation of the contained momenta of a given sub-current via

$$\epsilon(m_1, m_2) = (-1)^{\chi(m_1,m_2)}\,, \tag{4.8}$$

with

$$\chi(m_1, m_2) = \sum_{i=n}^{2} \hat{m}_{1,i} \left(\sum_{j=1}^{i-1} \hat{m}_{2,j} \right)\,, \tag{4.9}$$

where $\hat{m}_{i,j}$ is the j-th bit of the integer and hatted bits are set to 0 if the corresponding external particle is a boson. The Dyson-Schwinger equation can be written in diagrammatic form as



(4.10)

where the first and second term correspond to the first and second term in Eq. (4.6), respectively. In practical terms, if P is exactly one of the external momenta p_i then the current is simply given by the wave function of the external state, otherwise it is computed using the recursive combination of two (or more) sub-currents. The initial conditions for the external states are given by the well-known polarization vectors for bosons and spinors for fermions

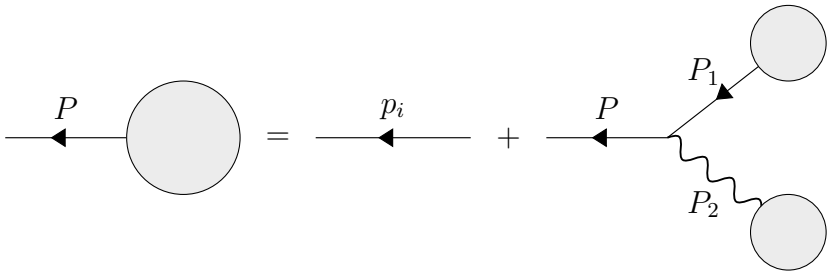
$$\begin{aligned} A^\mu(p_i) &= \epsilon^\mu_\lambda(p_i)\,, & \lambda &= \pm 1, 0 \\ \psi(p_i) &= \begin{cases} u_\lambda(p_i)\,, & \text{where } p_i^0 \geq 0 \\ v_\lambda(-p_i)\,, & \text{where } p_i^0 \leq 0 \end{cases} \\ \bar{\psi}(p_i) &= \begin{cases} \bar{u}_\lambda(p_i)\,, & \text{where } p_i^0 \geq 0 \\ \bar{v}_\lambda(-p_i)\,, & \text{where } p_i^0 \leq 0 \end{cases} \end{aligned} \tag{4.11}$$

The index λ is used to label the helicity states. Accordingly, for massless vectors, λ would take only two values, whereas for scalars only one value would be allowed. Whenever P does not correspond to an external particle, one instead checks all generated sub-currents with momenta P_1 and P_2, which fulfil $P = P_1 + P_2$, and applies the corresponding vertex function. The recursion can be implemented both from a bottom-up and top-down approach. The HELAC algorithm opts for the bottom-up approach. It is convenient to assign to each integer m_i in the binary representation a level

$$l = \sum_{i=1}^{n} \hat{m}_{i,j}\,. \tag{4.12}$$

External states have the level 1, currents with two momenta inside have the level 2 and the final current in the recursion is level $(n-1)$. Currents of a given level can only ever depend on currents of the lower levels, which implies a straightforward ordering of the algorithm. Level 1 can be trivially filled from initial conditions. Starting

from level 2 up to $(n-1)$, one simply searches for all combinations of lower-level currents consistent with any specific momentum configuration and combines them according to the Dyson-Schwinger equation. Continuing with the QED example, we still require the corresponding Dyson-Schwinger equation for the generation of fermion currents. The diagrammatic form is



, (4.13)

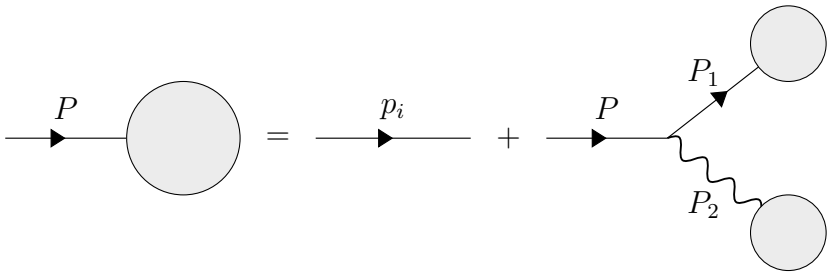
which is equivalent to

$$\psi(P) = \sum_{i=1}^{n} \delta_{P=p_i}\psi(p_i) + \sum_{P=P_1+P_2} (ie)\,\mathcal{P}\,\not{b}(P_2)\psi(P_1)\epsilon(P_1,P_2)\,, \tag{4.14}$$

where the fermion propagator $\mathcal{P}$ is given by

$$\mathcal{P} = \frac{i}{\not{P} - m}\,, \tag{4.15}$$

where m is the fermion mass. Similarly, currents corresponding to anti-fermions are given by



, (4.16)

or

$$\bar{\psi}(P) = \sum_{i=1}^{n} \delta_{P=p_i}\bar{\psi}(p_i) + \sum_{P=P_1+P_2} (ie)\,\bar{\psi}(P_1)\not{A}(P_2)\bar{\mathcal{P}}\,\epsilon(P_1,P_2) \tag{4.17}$$

with the propagator $\bar{\mathcal{P}}$ for an anti-fermion of mass m given by

$$\bar{\mathcal{P}} = \frac{-i}{\not{P} + m}\,. \tag{4.18}$$

Once the recursion has reached the final level $(n-1)$ one computes the scattering matrix element with the relation

$$\mathcal{M}(p_1,...,p_n) = \begin{cases} A_0^{\mu}(P_i)A_{\mu}(p_i) & i \in \{B\} \\ \bar{\psi}_0(P_i)\psi(p_i) & i \in \{F\} \\ \bar{\psi}(p_i)\psi_0(P_i) & i \in \{A\} \\ S_0(P_i)S(p_i) & i \in \{S\} \end{cases} \tag{4.19}$$

where $\{B\}$ is the set of bosons, $\{F\}(\{A\})$ the set of (anti-)fermions and $\{S\}$ the set of scalars and the endpoint momentum P_i is given by

$$P_i = \sum_{j \neq i} p_i \,. \tag{4.20}$$

The subscript 0 denotes the currents generated by iterating the recursive algorithm up to the endpoint current. The i-th particle is arbitrarily chosen as the endpoint of the recursion. The momentum of the endpoint is the only one that does not appear in the recursion. In HELAC the endpoint is always chosen to be the first particle, which implies that the integer representation of the momentum vector is always even. Other choices can be useful as well. For example, the gauge invariance of the amplitude can tested by replacing the wave function of a gluon with the respective momentum. If that gluon is chosen to be the endpoint of the recursion, then the evaluation of this test does not require any reevaluation of the recursive algorithm. Different helicity configurations of the external particles can be computed by simply adjusting the initial conditions. In total, one obtains $2^{l_2}3^{l_3}$ combinations, where l_2 and l_3 are the numbers of external particles with 2 and 3 helicity states, respectively. Performing the sum over helicities explicitly is very expensive, especially for computations with many final state particles. Thus, the HELAC-NLO framework offers the ability to treat this summation using Monte Carlo techniques instead. In principle, one can draw random integers and simply compute helicity states corresponding to the random labels. However, amplitudes with different helicity configurations can differ by orders of magnitude, which makes this approach numerically inefficient. Instead, it is useful to define a superposition [206]. For a massless vector it reads:

$$\epsilon^{\mu}_{\phi} = e^{i\phi}\epsilon^{\mu}_{+} + e^{-i\phi}\epsilon^{\mu}_{-} \,, \tag{4.21}$$

where the angle $\phi \in [0, 2\pi)$ is to be integrated over using Monte Carlo sampling. The expected summation is then recovered by the identity

$$\frac{1}{2\pi}\int_0^{2\pi} d\phi\, \epsilon^{\mu}_{\phi}(p)(\epsilon^{\nu}_{\phi}(p))^* = \sum_{\lambda=\pm 1} \epsilon^{\mu}_{\lambda}(p)(\epsilon^{\nu}_{\lambda}(p))^* \,. \tag{4.22}$$

For each phase-space point one selects and continues to optimize the angle ϕ and then computes just a single helicity amplitude using the superposition of the polarisation vectors ϵ_{ϕ}.

A naive extension to colored particles is straightforward as well. External particles can be similarly labelled in color space and the recursion repeated for each configuration of the external particles. Gluons enter in the adjoint representation, while quarks enter in the fundamental representation. Overall one obtains $(N_C^2-1)^{n_g}N_C^{n_q}$ different configurations, where n_g is the number of external gluons, n_q the number of external quarks and $N_C = 3$ the number of colors. Vertex functions similarly have to include the corresponding factors based on the color information of the currents entering as input and propagate this information to the output currents. However, the computational cost of such an approach would be very high. LHC processes often involve many jets and thus highly benefit from improved computational performance in the color algebra. More structured approaches to the color handling have been developed [207]. HELAC relies on the so-called color-flow decomposition representation, which is discussed in the following.

4.1.2 Color-flow decomposition

At its core, the color-flow decomposition avoids the usual decomposition of the gluon field in the adjoint representation $A_\mu = A_\mu^a t^a$, $a = 1, ..., 8$. Instead, one operates directly with an $N_C \times N_C$ matrix, $(A_\mu)_j^i$. Following closely Ref. [207] we write the Yang-Mills Lagrangian as

$$\mathcal{L} = -\frac{1}{4}(F^{\mu\nu})_j^i (F_{\mu\nu})_i^j + i\bar{\psi}_i\gamma^\mu \left(\delta_j^i \partial_\mu + i\frac{g}{\sqrt{2}}(A_\mu)_j^i\right)\psi^j - m\bar{\psi}_i\psi^i \,, \tag{4.23}$$

where the field-strength tensor takes the usual form

$$(F_{\mu\nu})_j^i = \partial_\mu (A_\nu)_j^i - \partial_\nu (A_\mu)_j^i + i\frac{g}{\sqrt{2}}\left[(A_\mu),(A_\nu)\right]_j^i \,. \tag{4.24}$$

Additional care has to be taken when summing over the (anti-)fundamental indices of external gluons. The gluon field only contains $N_C^2 - 1 = 8$ independent components, but N_C^2 index pairs enter the sum. The usual adjoint basis accounts for this by requiring a traceless generator basis $tr(t^a) = 0$. To achieve the same effect in the color-flow decomposition basis, one instead projects the external gluon states using the projector

$$P_{j,j'}^{i,i'} = \delta_{j'}^i \delta_j^{i'} - \frac{1}{N_C}\delta_j^i \delta_{j'}^{i'} \,. \tag{4.25}$$

The gluon propagator similarly decomposes into two terms

$$\left\langle (A_\mu)_{j_1}^{i_1} (A_\nu)_{j_2}^{i_2} \right\rangle \sim \delta_{j_2}^{i_1}\delta_{j_1}^{i_2} - \frac{1}{N_C}\delta_{j_1}^{i_1}\delta_{j_2}^{i_2} \,, \tag{4.26}$$

where the first one connects the colors across the propagator and the second one generates a $U(1)$-like contribution with both gluons self-connected. The Feynman rules for the vertices also become products of Kronecker-deltas in color space. For example, the two possible color-flow configurations of the 3-gluon vertex can be written as $\delta_{j_2}^{i_1}\delta_{j_3}^{i_2}\delta_{j_1}^{i_3}$ and $\delta_{j_3}^{i_1}\delta_{j_1}^{i_2}\delta_{j_2}^{i_3}$, where the indices label the colors of the 3 gluon fields. The fermion-gluon interaction vertex only allows for one color-flow given by $\delta_{j_1}^{i_q}\delta_{j_q}^{i_1}$. The four-gluon vertex generates the six non-cyclic permutations of $\delta_{j_2}^{i_1}\delta_{j_3}^{i_2}\delta_{j_4}^{i_3}\delta_{j_1}^{i_4}$. Rather than treating the propagator separately, HELAC always multiplies the numerator of the propagator into the current definition. The projector applied to external gluon states also has the same color factors as the propagator. This allows us to always multiply this structure into the $\bar{q}qg$ vertex, which is then described by two color flows instead of one [204, 208]. We will later see that both types of contributions can be extracted with a unified algorithm, which makes this formulation easy to apply in automation. The anti-cyclic structure of the three- and four-gluon vertices effectively prohibits any coupling to the aforementioned $U(1)$ gluon, which leaves the two and four respective color flows intact. With quarks and gluons on equal footing, one can decompose the matrix elements with $k = n_g + n_{q\bar{q}}$ gluons and quark-pairs as

$$\mathcal{M}_{j_1,...,j_k}^{i_1,...,i_k} = \sum_\sigma \delta_{j_1}^{i_{\sigma_1}}\delta_{j_2}^{i_{\sigma_2}}...\delta_{j_k}^{i_{\sigma_k}} A_\sigma = \sum_\sigma D_\sigma A_\sigma \,, \tag{4.27}$$

where A_σ are the color-stripped partial amplitudes and σ denotes a permutation of the $\{1, 2, .., k\}$ set. Notably each color-stripped amplitude is separately gauge-

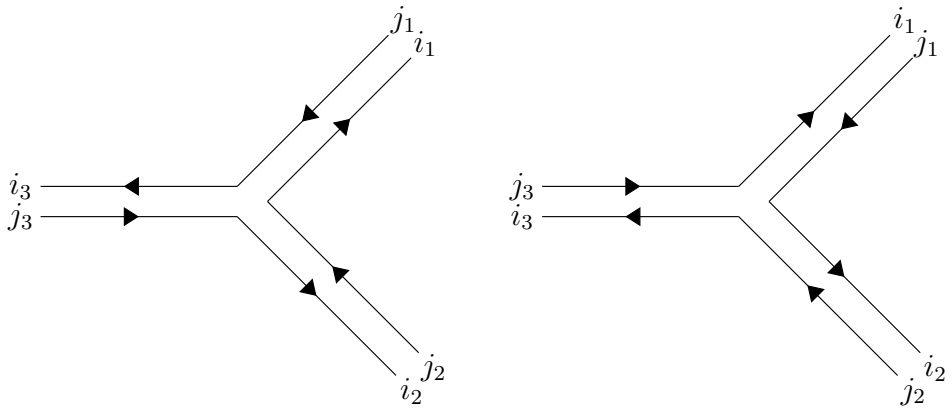


Figure 4.1: *Diagrammatic representation of the two contributions to the SM triple-gluon vertex in the color-flow decomposition. The color indices of gluons are labelled by i_k and j_k, where $k = 1, 2, 3$.*

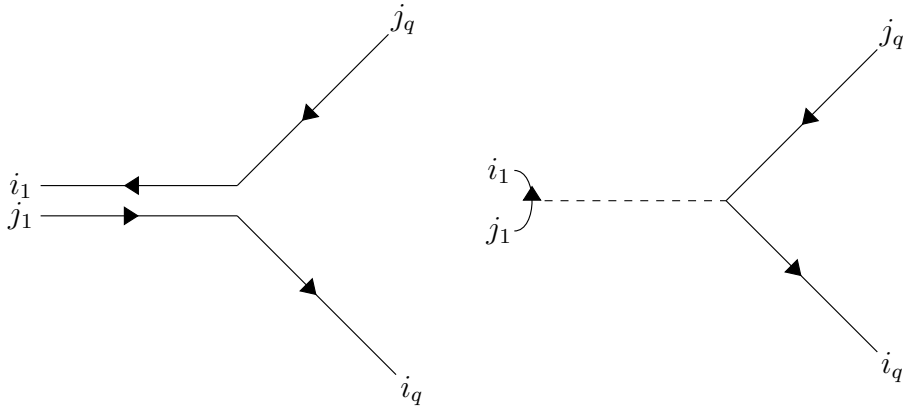


Figure 4.2: *Diagrammatic representation of the two contributions to the quark-gluon interaction vertex in the color-flow decomposition. The dashed line represent the $U(1)$ gluon contribution and carries a factor $1/N_C$. The color indices of the gluon are labelled with i_1 and j_1, while the quark pair carries i_q and j_q.*

invariant, which can be used to validate the implementation. Squared matrix elements are obtained with the relation

$$\sum_{\{i\},\{j\}} |\mathcal{M}^{i_1,\ldots,i_k}_{j_1,\ldots,j_k}|^2 = \sum_{\sigma,\sigma'} A^\dagger_\sigma C_{\sigma,\sigma'} A_{\sigma'} \,, \tag{4.28}$$

where $C_{\sigma,\sigma'}$ is the color matrix connecting two color configurations σ and σ'. It can be written as

$$C_{\sigma,\sigma'} = \sum_{\{i\},\{j\}} D_\sigma D^\dagger_{\sigma'} = N_C^{m(\sigma,\sigma')} \,, \tag{4.29}$$

where $1 \leq m(\sigma,\sigma') \leq k$ is the number of common cycles after multiplying the two products of Kronecker deltas [204]. In this formulation, one only has to compute the partial amplitudes of a given structure once and the corresponding factors in N_C are included through the color matrix. In practice, the derivation of vertices in the color-flow decomposition is not performed starting from the Lagrangian in Eq. (4.23). Instead, we exploit repeated use of Fierz-identities to translate the adjoint representation to the color-flow basis. The details are outlined separately in Section

5.1.1. One feature currently not exploited is the relations among several of the color-stripped amplitudes arising from the tracelessness of the gluons. For example, the following relation holds:

$$0 = \delta_{i_1}^{j_1} \mathcal{M}_{j_1,\ldots,j_k}^{i_1,\ldots,i_k} , \tag{4.30}$$

where i_1 and j_1 are the color indices of the gluon. Repeating the same procedure for all external gluons yields a system of equations that can be solved for the sub-leading partial amplitudes [207, 209]. In HELAC, we instead employ Monte Carlo techniques to reduce the computational cost of summing over many color configurations. Colors are assigned in the fundamental representation. The color-flow configurations compatible with the assigned colors are then evaluated. In this context, the number of partial amplitudes required is usually much smaller than the number needed to solve the equations. The trade-off comes in the form of the additional phase-space points that need to be evaluated to resolve the Monte Carlo integration in color space.

4.2 Virtual corrections

The calculation of virtual corrections is conceptionally more challenging than the tree-level matrix elements. At the 1-loop level, one four-momentum remains unconstrained and has to be integrated over. During the integration, two types of divergences arise. Ultra-violet (UV) divergences corresponding to the large momentum limit and infrared (IR) divergences corresponding to singularities at low loop momenta. In order to successfully perform the Monte Carlo integration over the phase-space of the external momenta one must first isolate the finite part of the integrals by introducing a suitable regularization procedure. In practice, dimensional regularization is particularly desirable. One performs an analytical continuation to $d = 4 - 2\varepsilon$ dimensions in the loop-momentum space. The obtained result can then be expanded in ε to obtain both the finite part and the singular pieces. The UV divergences can be analytically removed through the well-known renormalization procedure by introducing a set of local counterterms. The remaining IR divergences cancel against equivalent singularities in the real-emission corrections. While such a cancellation can be explicitly performed analytically, the numerical implementation is more difficult. Modern NLO Monte Carlo codes extract the singular behavior through semi-analytical means. HELAC-1LOOP employs the OPP method developed by **O**ssola, **P**apadopoulos and **P**ittau [201, 210] in 2006.

The OPP method is a technique to take an arbitrary 1-loop amplitude and perform a numerical reduction to scalar master integrals on the level of the integrand. These master integrals can then be evaluated from precomputed analytical solutions. As a result, the technique combines the fast evaluation time of the numerical reduction with the stability of analytical solutions. Furthermore, the method requires very little knowledge about the structure of the amplitude, which makes it particularly well suited to processes with many final state particles and high numbers of massive propagators.

Like previously mentioned, we work in dimensional regularization and denote a d-dimensional object $\bar{q}$ with a bar and decompose it into a four-dimensional part q

and a $d-4$-dimensional part $\tilde{q}$ as $\bar{q}^2 = q^2 + \tilde{q}^2$. A general 1-loop amplitude may then be written as

$$A = \mu^{4-d} \sum_s \int \frac{d^d\bar{q}}{(2\pi)^d} \frac{\bar{N}_s(\bar{q})}{\prod_{i=0}^{m_s-1} \bar{D}_{s_i}(\bar{q})}\,, \tag{4.31}$$

where $\bar{q}$ is the loop-momentum, μ is the renormalization scale, s runs over the full set of loop topologies, $\bar{N}_s$ is the numerator function for a given topology [1] and an m_s-point function is accompanied by denominators with mass m_{s_i} given by

$$\bar{D}_{s_i} = (\bar{q} + p_{s_i})^2 - m_{s_i}^2\,. \tag{4.32}$$

The momenta p_{s_i} are functions of the external, four-dimensional momenta, which enter the loop. It is understood that the numerator function can contain contributions from several Feynman diagrams at once. Only the structure of the denominator is needed to characterize the loop topology. Topologies can of course differ by loop assignment, but also through different propagator masses, flavors and color assignments. In principle, one can relax this condition slightly and reduce all contributions with identical propagator masses simultaneously. However, HELAC-NLO currently is not set up for this optimization.

4.2.1 Construction of 1-loop topologies

The loop topologies are constructed using a similar language as the HELAC algorithm itself. The starting point is an n-particle process and a given color configuration of the external particles. For the given process we need to generate all N-point functions with $N \leq n$. For one N at a time, we first classify the tree-like currents positioned at the corners of the loop. This is done using the same language as the original HELAC algorithm itself. An amplitude can be identified by its binary representation $2^n - 1$. The code now constructs all sets of binary integers $\{i_j | j = 1, ..., N\}$ such that

$$\sum_{j=1}^{N} i_j = 2^n - 1\,. \tag{4.33}$$

To avoid double-counting, equivalent topologies have to excluded. For that one can start with one starting topology, $\{i_1, ..., i_N\}$, ensuring that i_1 is the one odd integer. Then one generates all permutations of the remaining integers, which ensures that cyclical permutations are excluded. In addition, the configuration with i_j in reversed order compared to the starting point has to be excluded, e.g. for the starting configuration $\{i_1, i_2, i_3, i_4\}$ one should not include $\{i_1, i_4, i_3, i_2\}$. One can confirm by looking at the left side of Fig. 4.3 that such a reversed order would simply generate the same topology. The next phase of the algorithm dresses the chosen loop topology with particle flavors and colors. It is sufficient to simply pick one of the propagators in the loop, assign a flavor and two color indices, which represent the external colors that the propagator is connected to. Consistent with HELAC one assigns one of the indices to 0 for quarks. An extension to the EW sector would simply assign both indices to 0. The restriction on the flavor assures that a unique set of propagator masses is selected. Similarly, the restriction on the color is of technical nature, as the HELAC algorithm is currently not equipped to deal with currents of similar

[1] Loop-momentum independent parts of calculation may be included in the definition of $\bar{N}_s$.

momenta with different color connections simultaneously. Just like the LO recursion, we thus compute these contributions sequentially on their own. In the next step, one propagator line will be replaced by two additional external particles, which makes it easy to understand why at most two additional color lines are needed. From here, one can apply the regular HELAC recursion to generate the numerator function, which will only generate non-zero results if a valid configuration was chosen. In practice, it is useful to also assign the color connections and flavors for the remaining propagators in the loop. Using the color representation and flavor information of the loop propagators, one can restrict the actual number of topologies that need to be sent to the recursion by ensuring color conservation. In the current implementation this is faster than using the LO recursion to test for valid configurations. However, the OPP reduction only requires a consistent set of propagator masses. A modified version of the tree-level recursion could easily support the simultaneous generation of all loop topologies of a given set of propagator masses. Doing so would substantially accelerate the evaluation of loop amplitudes, because many topologies would be combined into one numerator function and consequently could be treated with a single reduction.

4.2.2 Numerical computation of the numerator function

Before detailing the reduction itself, we discuss how to obtain the numerator function numerically. In Eq. (4.31) the numerator function is given as a d-dimensional object. One can use recursive techniques conceptually identical to the HELAC algorithm to obtain the four-dimensional part only. The mismatch is treated separately by introducing so-called rational parts. We postpone this discussion and focus on the four-dimensional part for now. Starting with one of the aforementioned loop topologies, one cuts one of the lines as demonstrated in Figure 4.3. An $(n+2)$-particle process is obtained. The numerator of the cut propagator line is recovered

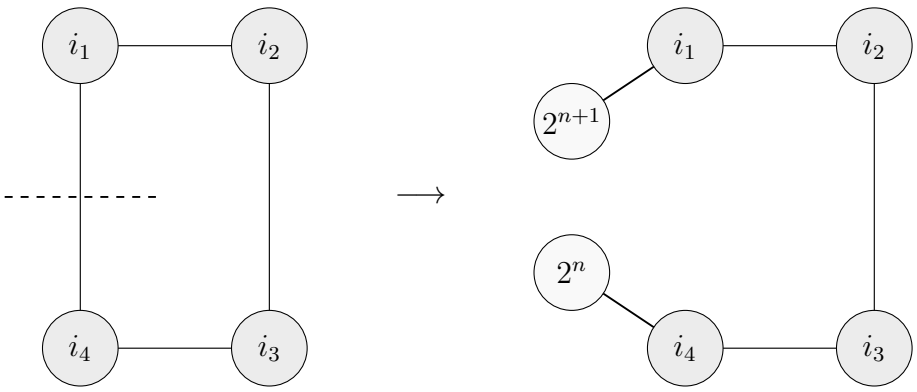


Figure 4.3: *The numerator of the box topology can be calculated by cutting one line and computing the HELAC recursion with two additional particles. The blobs i_j can contain multiple external momenta following the HELAC binary representation. Consequently, multiple structures can be contained within.*

through the sum over the helicity states of the cut line. We work in the Feynman gauge for this purpose, but similar considerations apply for any renormalizable gauge. Because the scalar propagator is trivial one can set the wave functions of the

extra particles to

$$\varepsilon = (1, 0, 0, 0)\,. \tag{4.34}$$

For on-shell fermions one can apply the completeness relations

$$\sum_{s=1}^{2} u_{(s)}(q)\bar{u}_{(s)}(q) = \not{q} + m \qquad \text{and} \qquad \sum_{s=1}^{2} v_{(s)}(q)\bar{v}_{(s)}(q) = \not{q} - m\,, \tag{4.35}$$

where m is the mass of the fermion line [2]. However, propagators will generally be off-shell in the loop and two additional auxiliary states are required in the sum to recover the correct expression. Similarly, the numerator of the vector-boson propagator in the Feynman gauge is given by $-g^{\mu\nu}$. However, considering the completeness relation over the two physical states of the massless boson yields

$$\sum_{\pm 1} \varepsilon^{\mu}_{(\lambda)}(q)(\varepsilon^{\nu}_{(\lambda)}(q))^* = -g^{\mu\nu} + \frac{q^\mu n^\nu + q^\nu n^\mu}{q\cdot n} - \frac{n^2 q^\mu q^\nu}{(q\cdot n)^2}\,, \tag{4.36}$$

where n is an auxiliary four vector. To decompose purely the propagator piece, one instead needs to sum over four states, with two of them being unphysical states. A simple choice for these can be given using the light-cone representation of the four-vector $V^A = (V^0 + V^3, V^0 - V^3, V_1 + iV^2, V_1 - iV^2)^T$. We write the down four polarisation states

$$\begin{aligned}
\varepsilon_1 &= (1,0,0,0)^T\,, & \bar{\varepsilon}_1 &= (0,1,0,0)^T \\
\varepsilon_2 &= (0,1,0,0)^T\,, & \bar{\varepsilon}_2 &= (1,0,0,0)^T \\
\varepsilon_3 &= (0,0,1,0)^T\,, & \bar{\varepsilon}_3 &= (0,0,0,-1)^T \\
\varepsilon_4 &= (0,0,0,1)^T\,, & \bar{\varepsilon}_4 &= (0,0,-1,0)^T\,,
\end{aligned} \tag{4.37}$$

and the decomposition takes the form

$$\sum_{\lambda=0}^{4} \varepsilon^{A}_{(\lambda)}(q)\bar{\varepsilon}^{B}_{(\lambda)}(q) = -2g^{AB}\,, \tag{4.38}$$

where g^{AB} is the equivalent of the metric tensor in the light-cone representation. A similar argument holds for the case of massive gauge bosons. In practical terms, one simply prepares the two additional external wave functions, while carefully constraining flavors, color and the helicity state to account for the corresponding numerator in the 1-loop topology. The $(n+2)$ tree-level process can then be computed using the regular HELAC algorithm and the resulting amplitudes summed over the helicities of the extra particles to obtain the full numerator. It is clear that reducing the number of states of the additional particles increases the computational performance. Splitting a scalar line will always result in the fastest evaluation time, while fermions can still be optimized to only require a sum over two states in specific circumstances. Vector bosons are always the slowest in practice. HELAC-1LOOP does not implement such an optimization at present. With the four-dimensional numerator computed, we return to the 1-loop amplitude reduction.

[2] This is indeed very common in the OPP algorithm as we will understand in the following sections.

4.2.3 Reduction of loop amplitudes

At the core of the OPP method lies the amplitude decomposition in the $d \to 4$ limit as

$$A = \sum_i d_i \, \mathrm{Box_i} + \sum_\mathrm{i} \mathrm{c_i} \, \mathrm{Triangle_i} + \sum_\mathrm{i} \mathrm{b_i} \, \mathrm{Bubble_i} + \sum_\mathrm{i} \mathrm{a_i} \, \mathrm{Tadpole_i} + \mathrm{R} \,, \tag{4.39}$$

where $R = R_1 + R_2$ is the rational contribution of first and second kind, and Box, Triangle, Bubble and Tadpole correspond to the scalar 1-loop function basis. The R_2 terms capture the contribution of the ε-dimensional part of the numerator function and can be computed by inserting counterterm-like contributions in tree-level amplitudes. To establish that Eq. (4.39) indeed holds, we first study the scalar 5-point function. In the next step, we extend the argument to integrals of higher rank.

In d-dimensions, $(d+m)$-point functions can be reduced to a basis of at most d-point functions, which can be proven through various means. We will follow the proof provided in Ref. [211]. The denominators are defined as in Eq. (4.32), but the loop-momentum is shifted such that the first propagator is just given by $\bar{q}^2 - m_0^2$. We further introduce notation for the contraction of an ϵ tensor with four-vectors

$$\epsilon^{p_1 p_2 p_3 p_4} = \epsilon^{\mu\nu\rho\sigma} p_{1\mu} p_{2\nu} p_{3\rho} p_{4\sigma} \,. \tag{4.40}$$

The integral

$$\int \frac{d^4 q \, q_\mu}{D_0 D_1 D_2 D_3} = c_1 p_{1\mu} + c_2 p_{2\mu} + c_3 p_{3\mu} \,, \qquad c_i = \text{constant} \tag{4.41}$$

can only depend on linear functions of the external four-momenta, which we denote by p_i^μ with $i = 1, 2, 3$. Contracting with the fully anti-symmetric tensor thus yields

$$\int \frac{d^4 q \, \epsilon^{p_1 p_2 p_3 q}}{D_0 D_1 D_2 D_3} = 0 \,. \tag{4.42}$$

In four-dimensions one can find at most 4 linearly independent vectors. As a consequence, one can find the Schouten identity, which implies that the sum of 5-vectors contracted with the Levi-Cevita tensor vanishes [212]. The proof of the reduction starts from

$$\epsilon^{p_1 p_2 p_3 p_4} q_\mu = \epsilon^{\mu p_2 p_3 p_4} q \cdot p_1 + \epsilon^{p_1 \mu p_3 p_4} q \cdot p_2 + \epsilon^{p_1 p_2 \mu p_4} q \cdot p_3 + \epsilon^{p_1 p_2 p_3 \mu} q \cdot p_4 \,. \tag{4.43}$$

It is useful to define basis vectors corresponding to the contraction of the Levi-Cevita tensors

$$\begin{aligned} v_1^\mu &= \epsilon^{\mu p_2 p_3 p_4}/a \,, & v_2^\mu &= \epsilon^{p_1 \mu p_3 p_4}/a \,, \\ v_3^\mu &= \epsilon^{p_1 p_2 \mu p_4}/a \,, & v_4^\mu &= \epsilon^{p_1 p_2 p_3 \mu}/a \,, \end{aligned} \tag{4.44}$$

where $a = \epsilon^{p_1 p_2 p_3 p_4}$ and $\bar{a} = \epsilon_{p_1 p_2 p_3 p_4} = -a$. This basis is called van Neerven - Vermaseren basis and is not necessarily orthogonal. For integrals with less than four independent vectors we can still define it, but must introduce additional transverse vectors to span the full space. The product $a\bar{a}$ is the Gram-determinant of the system. Contracting Eq. (4.43) with q^μ and simplifying it using the basis vectors yields

$$a q^2 = \sum_{i=1}^{4} a (q \cdot v_i)(q \cdot p_i) \,, \tag{4.45}$$

which can be further re-written as

$$2a(D_0 + m_0^2) = a\sum_{i=1}^{4}(q \cdot v_i)(D_i - D_0 - r_i)\,, \tag{4.46}$$

after substituting $q^2 = D_0 + m_0^2$ and $q \cdot p_i = \frac{1}{2}(D_i - D_0 - r_i)$, with $r_i = p_i^2 - m_i^2 + m_0^2$. Dividing the whole expression by $\prod_{i=0}^{4} D_i$ will recover the desired 5-point function.

The contributions proportional to $(q \cdot v_i)D_i$ vanish under integration due to Eq. (4.42) after cancelling the D_i in the numerator against one of the ones in the denominator and shifting the loop-momentum. After reordering the remaining terms we can write

$$0 = \int d^4q \frac{2am_0^2 + \left(2a + a\sum_{i=1}^{4}(q \cdot v_i)\right) D_0 + a\sum_{i=1}^{4} r_i(q \cdot v_i)}{D_0 D_1 D_2 D_3 D_4}\,, \tag{4.47}$$

where the first term is proportional to the 5-point scalar function. The contribution of the type $\sum_{i=1}^{4}(q \cdot v_i)D_0/(D_0 D_1 D_2 D_3 D_4)$ can be further simplified. We substitute

$$a\sum_{i=1}^{4}(q \cdot v_i) = \epsilon^{k(p_2-p_1)(p_3-p_1)(p_4-p_1)} - a\,, \tag{4.48}$$

where the loop-momentum was shifted according to $k = q - p_1$. Under the integral one can then eliminate the first term according to Eq. (4.42). Inserting the result back into Eq. (4.47), we obtain

$$0 = \int d^4q \frac{aD_0 + 2am_0^2 + \sum_{i=1}^{4} r_i a(q \cdot v_i)}{D_0 D_1 D_2 D_3 D_4}\,. \tag{4.49}$$

The last term can be further simplified by multiplying with $\bar{a}$ and introducing the auxiliary vector $w = a\sum_{i=1}^{4} r_i v_i$. One finds

$$a\bar{a}q \cdot v_1 = \epsilon_{p_1 p_2 p_3 p_4}\epsilon^{q p_2 p_3 p_4} = a^2\sum_{i=1}^{4}(q \cdot p_i)(v_1 \cdot v_i)\,, \tag{4.50}$$

and

$$a\sum_{i=1}^{4} r_i\bar{a}(q \cdot v_i) = a\sum_{i=1}^{4}(q \cdot p_i)(v_i \cdot w) = \frac{a}{2}\sum_{i=1}^{4}(D_i - D_0 - r_i)(v_i \cdot w)\,, \tag{4.51}$$

where $(q \cdot p_i)$ was replaced using the definitions of the propagators. Finally inserting the last identity back into the integral expression we find

$$0 = \int d^4q \frac{a\bar{a}(2m_0^2 - \frac{w^2}{2}) + D_0 a\bar{a} - \frac{a}{2}D_0\left(\sum_{i=1}^{4}(v_i \cdot w)\right) + \frac{a}{2}\left(\sum_{i=1}^{4} D_i(v_i \cdot w)\right)}{D_0 D_1 D_2 D_3 D_4}\,. \tag{4.52}$$

After cancelling all the propagators and defining the scalar 5-point function as $\mathcal{I}^{(5)}_{01234}$ we obtain the final decomposition

$$\begin{aligned}\mathcal{I}^{(5)}_{01234} = \frac{1}{w^2 - 4a\bar{a}m_0^2}\Bigg(& D_{1234}\left[2a\bar{a} - a\sum_{i=1}^{4} w \cdot v_i\right] + D_{0234}(av_1 \cdot w) \\ &+ D_{0134}(av_2 \cdot w) + D_{0124}(av_3 \cdot w) + D_{0123}(av_4 \cdot w)\Bigg)\,,\end{aligned} \tag{4.53}$$

where $D_{i_0 i_1 i_2 i_3}$ is the four-point scalar integral with denominators i_0, i_1, i_2 and i_3. The case for six-point functions and higher is much simpler due to the presence of additional external momenta. Multiplying Eq. (4.43) with the additional four-vector p_5 yields

$$a(q \cdot p_5) = a \sum_{i=1}^{4} (v_i \cdot p_5)(q \cdot p_i) \,, \tag{4.54}$$

which is only a linear function in q instead of the quadratic one we encountered in Eq. (4.45). The equivalent of Eq. (4.52) is then given by

$$0 = a \int \frac{\sum_{i=1}^{4} r_i (v_i \cdot p_5) + (D_5 - r_5) + D_0 (\sum_{i=1}^{4} p_5 \cdot v_i - 1) - \sum_{i=1}^{4} D_i (v_i \cdot p_5)}{\prod_{i=0}^{5} D_i} d^4 q \,, \tag{4.55}$$

which yields a decomposition of the six-point function into multiple five-point functions on the level of the integrand already. This result can easily be extended to higher-point functions from this point, by simply adding more propagators to the denominator. This is in contrast to the 5-point case, where explicit integration was needed to remove certain terms.

The case of higher rank-m integrals, that is, integrals with m additional loop-momenta in the numerator

$$\mathcal{I}_N = \int d^d \bar{q} \frac{\prod_{i=1}^{m} \bar{q} \cdot u_i}{\prod_{i=0}^{N-1} \bar{D}_i} \,, \tag{4.56}$$

where the u_i are functions of the external momenta, can be reduced to purely scalar integrals with at most N-denominators as well as 4-point tensor functions [213]. Following the argument for the 5-point scalar function one can then further reduce the N-point functions to 4-point functions as well. To further understand the reduction of tensor-integrals we follow the procedure present in Ref. [214]. First we focus on the case $N > 4$. In this case, we can formulate the van Neerven - Vermaseren basis, which fully spans the vector space. The d-dimensional loop momentum can be decomposed as

$$\bar{q}^\mu = \sum_{i=1}^{4} (\bar{q} \cdot p_i) v_i^\mu + (\bar{q} \cdot n_\varepsilon) n_\varepsilon^\mu \,, \tag{4.57}$$

where n_ε is an additional basis vector living purely in the $(d-4)$-dimensional space. It is transverse to the four-dimensional part of any four-vector, i.e. $n_\varepsilon \cdot p = 0$. The general strategy then involves the replacement of one $u_i \cdot \bar{q}$ term in the numerator by inserting the decomposition of the loop-momentum. By rewriting the scalar product $\bar{q} \cdot p_i$ with the propagator definition, we obtain

$$\bar{N}(\bar{q}) = \frac{1}{2} \prod_{i=1}^{m-1} (\bar{q} \cdot u_i) \left[\left(\sum_{j=1}^{4} u_m \cdot v_j \right) \left(\bar{D}_j - \bar{D}_0 - \left(p_j^2 - m_j^2 + m_0^2 \right) \right) \right] \,, \tag{4.58}$$

where we have used that the physical vector has no $(d-4)$-dimensional component, i.e. $u_m \cdot n_\varepsilon = 0$. The first two terms in the expansion of the numerator will generate $(N-1)$-point functions with rank $(m-1)$ after cancelling the propagators against the denominator. The last term yields another N-point function, but with the rank reduced to $(m-1)$. Clearly, we can iterate this strategy and reduce the N-point function to a scalar one. The generated $(N-1)$-point contributions can also be

reduced using the same strategy as long as the number of propagators is above four. Together with the previous discussion on the reduction of the N-point scalar integrals for $N > 4$, we are only left with integrals with at most 4 propagators in the denominator. The reduction of tensor integrals with $N \leq 4$ denominators can be performed as follows. We construct the basis from the external momenta p_i, but now with additional basis vectors n_i transverse to our physical momenta. This ensures that the entire vector space can be spanned. The loop-momentum is then decomposed as

$$\bar{q}^\mu = \sum_{i=1}^{N-1} (\bar{q} \cdot p_i) v_i^\mu + \sum_{i=N}^{4} (\bar{q} \cdot n_i) n_i^\mu + (\bar{q} \cdot n_\varepsilon) n_\varepsilon^\mu \,. \tag{4.59}$$

As before, one can insert this definition into the numerator expression and replace products of the type $\bar{q} \cdot p_i$ with the help of the propagator definition. The resulting decomposition will look similar to the Eq. (4.58), but will include additional terms generated by the transverse vector contributions [3]. In the case of the rank-4 four-point function this term reads

$$\left(\prod_{i=1}^{3} u_i \cdot \bar{q} \right) (\bar{q} \cdot n_4)(u_4 \cdot n_4) \,, \tag{4.60}$$

which appears to be another rank-4 four-point contribution. To actually reduce this part one can repeat the same steps, which leads a sum of reduced terms as well as one rank-4 term proportional to

$$\left(\prod_{i=1}^{2} u_i \cdot \bar{q} \right) (\bar{q} \cdot n_4)^2 \,, \tag{4.61}$$

which can now be further rewritten using denominator definitions. One can further replace the product $(\bar{q} \cdot n_4)^2$, which generates terms which are either reduced in rank or in the number of denominators as well as one term proportional to the n_ε

$$(\bar{q} \cdot n_4)^2 = -(\bar{q} \cdot n_\varepsilon)^2 + \mathcal{O}(\bar{D}_i) + \text{constants} \,. \tag{4.62}$$

From here, we can simply repeat this replacement to generate at most one contraction of the loop momentum with the transverse vector and additional ones with the vector n_ε to some power. In four dimensions, i.e. when we don't include the n_ε part, all the 'spurious' loop-momentum dependent pieces will integrate to 0. This can be seen by writing the tensor integral in terms of the available Lorentz tensors, which will always yield a contraction of the type $p_i \cdot n_4 = 0$. We can apply a similar strategy to triangle, bubble and tadpole integrals. First, we introduce the van Neerven - Vermaseren basis, expressing the loop momentum in terms of the basis vectors. Then, we recursively replace squared terms in $n_i \cdot \bar{q}$ using the denominator definitions. Terms linear in any n_i will always vanish after integration for the same reason as in the 4-point function. For terms which are quadratic, one can always construct structures of the type $(q \cdot n_i)^2 - (q \cdot n_j)^2$, which will also vanish under integration. Terms proportional to n_ε require additional care, as these contribute so-called rational terms, which do not vanish in general. We delay this discussion

[3] Terms like $n_\varepsilon \cdot u_i$ drop out, because u_i are functions of physical vectors only.

and for now focus on the four-dimensional part of the numerators. We combine our finding for the reduction of the non-zero ranked integrals with the result for the reduction of the scalar N-point function with $N > 4$ [4]. When working in 4 dimensions, the final decomposition will only contain the scalar integrals with 1, 2, 3 and 4 denominators.

4.2.4 The OPP method

With the gained intuition on the reduction of 1-loop amplitude we can return to the OPP method. The OPP method as implemented in CUTTOOLS is designed to be used together with numerical approaches to obtain the numerators function. Rather than performing the reduction in d dimensions one instead performs it in four dimensions and obtains additional terms generated by the analytic continuation separately. Following the implementation presented in Ref. [215], the now four-dimensional numerator can be written in terms of four-dimensional propagators D_{s_i}

$$\begin{aligned} N_s(q) = & \sum_{i_0<i_1<i_2<i_3}^{m_s-1} \Big[d(i_0,i_1,i_2,i_3) + \tilde{d}(q;i_0,i_1,i_2,i_3)\Big] \prod_{i\neq i_0,i_1,i_2,i_3}^{m_s-1} D_{s_i} \\ & + \sum_{i_0<i_1<i_2}^{m_s-1} [c(i_0,i_1,i_2) + \tilde{c}(q;i_0,i_1,i_2)] \prod_{i\neq i_0,i_1,i_2}^{m_s-1} D_{s_i} \\ & + \sum_{i_0<i_1}^{m_s-1} \Big[b(i_0,i_1) + \tilde{b}(q;i_0,i_1)\Big] \prod_{i\neq i_0,i_1}^{m_s-1} D_{s_i} \\ & + \sum_{i_0}^{m_s-1} [a(i_0) + \tilde{a}(q;i_0)] \prod_{i\neq i_0}^{m_s-1} D_{s_i} \\ & + \tilde{P}(q) \prod_{i}^{m_s-1} D_{s_i}\,, \end{aligned} \tag{4.63}$$

where the a, b, c and d are the same coefficients as in Eq. (4.39) and the $\tilde{a}, \tilde{b}, \tilde{c}, \tilde{d}$ and $\tilde{P}$ are spurious terms that depend on the loop-momentum q. For brevity we suppress an additional index s in the decomposition coefficients, but note that the coefficients of course differ for each loop topology. The spurious terms cancel after integration as previously discussed. The ansatz simply decomposes the numerator according to its pole structure. The solution strategy simply involves determining all the coefficients in Eq. (4.63) by sampling different values of the loop momentum q. In general, this would involve inverting a rather large matrix, which becomes numerically difficult. Instead, a recursive approach is adopted. First for one specific configuration of $\{i_0, i_1, i_2, i_3\}$, one finds the two solutions $q_\pm$ for the loop momentum at which the four corresponding D_{i_j} vanish. Inserting the solution into Eq. (4.63) we obtain

$$N_s(q_\pm) = \Big[d(i_0,i_1,i_2,i_3) + \tilde{d}(q_\pm;i_0,i_1,i_2,i_3)\Big] \prod_{i\neq i_0,i_1,i_2,i_3}^{m_s-1} D_{s_i}(q_\pm)\,, \tag{4.64}$$

[4] The latter proof was performed in $D = 4$ dimensions. Before performing the limit $D \to 4$ one additional 5-point master integral can appear, which is needed to compute the expression of the coefficients for the lower point functions [214].

which can be used with the known functional form of the spurious term [5] to directly solve for the $d(i_0, i_1, i_2, i_3)$ and $\tilde{d}(i_0, i_1, i_2, i_3)$ coefficients. Choosing different configurations of $\{i_0, i_1, i_2, i_3\}$ one can extract all box coefficients. Next, one subtracts all boxes on both sides of the ansatz and repeats the procedure for all sets $\{i_0, i_1, i_2\}$ so that three denominators vanish. Keeping in mind that the spurious terms are decomposed into multiple terms [6], we then sample the numerator at sufficient number of points to extract all c and $\tilde{c}$ coefficients. One simply repeats this procedure for the b and a coefficients until the entire numerator is decomposed.

The procedure so far is designed to work in 4 dimensions. However, in typical LHC computations divergent diagrams might arise. These are regularized using dimensional regularization, that is, by an analytical continuation into d dimensions. Two types of additional rational terms need to be included as a result. The first kind, named R_1, arises due to the mismatch of the d-dimensional denominators in the general expression of the regularized one-loop amplitude and four-dimensional denominators in Eq. (4.63). These are generally process-dependent quantities and their computation is included in the CutTools program. One can generally obtain these by similarly sampling the loop momentum q. Consider how the d- and four-dimensional propagators are related:

$$\bar{D}_i = (\bar{q} + p_i)^2 - m_i^2 = \tilde{q}^2 + (q + p_i)^2 - m_i^2 = \tilde{q}^2 + D_i \,, \tag{4.65}$$

where we have used that $\tilde{q} \cdot p_i = 0$ for physical vectors p_i. In other words, we can shift the masses $m_i^2 \to m_i^2 - \tilde{q}^2$ in the formula for the four-dimensional numerator to reintroduce these terms [215]. The R_1 term is thus defined by

$$R_1 \equiv \frac{1}{(2\pi)^4} \int d^d\bar{q} \frac{f(\tilde{q}^2, q)}{\prod_{i=0}^{m-1} D_i} \,, \tag{4.66}$$

where the form of the function f is obtained from the four-dimensional expression of the numerator with the aforementioned shift of the masses. For at most rank-N N-point functions one ends up with replacements

$$\begin{aligned} b(ij) &\to b(ij) + \tilde{q}^2 b^{(2)}(ij) \,, \\ c(ijk) &\to c(ijk) + \tilde{q}^2 c^{(2)}(ijk) \,, \\ d(ijkl) &\to d(ijkl) + \tilde{q}^2 d^{(2)}(ijkl) + \tilde{q}^4 d^{(4)}(ijkl) \,. \end{aligned} \tag{4.67}$$

The additional coefficients can then be fitted by evaluating the numerator function again at several values of $\tilde{q}^2$.

The second kind of rational terms R_2 arises due to the d-dimensional components of the numerator function, which we have not included in the ansatz in Eq. (4.63). The general origin was already discussed in the description of the explicit reduction of higher rank integrals. Splitting off the four-dimensional part of the $(4 - 2\varepsilon)$-dimensional numerator function

$$\bar{N}(\bar{q}) = N(q) + \tilde{N}(\tilde{q}^2, q, \varepsilon) \,, \tag{4.68}$$

[5] The explicit form of the spurious terms was provided in Ref. [201] for integrals in renormalizable gauges.

[6] For renormalizable theories we would have to sample 7 values of the loop-momentum q. The decomposition of the spurious terms generally depends on the rank of the integral. Modern versions of CutTools implement the decomposition up to rank $n+1$, where n is the number of propagators in the loop.

leads to the following definition for the R_2 term

$$R_2 = \frac{1}{(2\pi)^4}\int d^d\bar{q}\,\frac{\tilde{N}(\tilde{q}^2,q,\varepsilon)}{\prod_{i=0}^{m-1}\bar{D}_i} \equiv \frac{1}{(2\pi)^4}\int d^d\bar{q}\,\mathcal{R}_2\,, \tag{4.69}$$

where we again stress that the d-dimensional denominators $\bar{D}_i$ are considered.

These terms are process independent and can be computed in a similar vein as the renormalization counterterms. One considers all Green's functions that are divergent and thus require dimensional regularization. One then computes the numerator function and splits off the four-dimensional part according to Eq. (4.68). This results in a finite number of integrals to be solved and to be implemented as counterterm vertices. For renormalizable theories only three additional integrals appear [201, 215]. These are given by

$$\begin{aligned}
\int d^d\bar{q}\frac{\tilde{q}^2}{\bar{D}_i\bar{D}_j} &= -\frac{i\pi}{2}\left[m_i^2+m_j^2-\frac{(p_i-p_j)^2}{3}\right]+\mathcal{O}(\varepsilon)\,,\\
\int d^d\bar{q}\frac{\tilde{q}^2}{\bar{D}_i\bar{D}_j\bar{D}_k} &= -\frac{i\pi^2}{2}+\mathcal{O}(\varepsilon)\,,\\
\int d^d\bar{q}\frac{\tilde{q}^4}{\bar{D}_i\bar{D}_j\bar{D}_k\bar{D}_l} &= -\frac{i\pi^2}{6}+\mathcal{O}(\varepsilon)\,.
\end{aligned} \tag{4.70}$$

The technique works the same for non-renormalizable theories. However, the ansatz for the spurious terms and the computation of the R_1 contribution have to be expanded to account for integrals, whose rank exceeds the number of propagators. Furthermore, additional R_2 have to be provided. For example, a dimension-6 Lagrangian needs to include R_2 terms with up to six legs as well.

To summarize the strategy, we compute the numerator function recursively using $(n+2)$-particle amplitudes. This step can be performed using the same tree-level algorithm that is used in HELAC with minimal modifications. Then, by sampling this numerator at several values of the loop momentum q we recursively compute the coefficients of the scalar integrals appearing in the reduction formula for the 1-loop function. The scalar loop integrals are then provided by the ONELOOP library. Similarly, R_1 terms are computed by sampling the numerator function with shifted masses. Finally, the R_2 term is implemented by dedicated tree-like vertices that can be computed for a given theory once and for all. Combining these 3 contributions with the regular UV counterterms provides a method to compute the renormalized 1-loop amplitude for arbitrary processes automatically.

4.2.5 Reweighting techniques for virtual corrections

The OPP method is vital for the automated calculation of 1-loop corrections. However, the reduction process involves a numerical fit using the four-dimensional numerator function. In practice, complicated processes can require $\mathcal{O}(100)$ evaluations for each separate loop topology. This is further amplified by the color and helicity degrees of freedom of the external states. The reduction acts on the level of the amplitude, which allows for optimizations similar to the tree-level case for color and helicity. However, even in this case one can quickly end up with hundreds of loop topologies, which each require many evaluations. Consequently, the evaluation of 1-loop matrix elements is very expensive compared to the tree-level case.

HELAC-1LOOP is fully capable of evaluating matrix elements in this way, which allows for direct study of individual phase-space points. This is useful for many cross-checks such as pole-cancellation, gauge-invariance checks as well as comparison with other programs. However, the computation of the virtual contribution to the cross sections would be very expensive. In practical calculations, one has to evaluate many phase-space points to optimize the phase-space generation. The pressure is further increased when helicity and color are treated using Monte Carlo methods as well. HELAC-1LOOP avoids this by employing a reweighting technique [127, 216, 217]. Note first that the virtual corrections are given by an interference term

$$\sigma_{ab}^{\rm LO+V} = \int dx_1 dx_2 d\phi_m f_a(x_1) f_b(x_2) (|\mathcal{M}|^2 + \mathcal{M}^*\mathcal{L} + \mathcal{M}^*\mathcal{L}) \tag{4.71}$$

where $\mathcal{M}$ and $\mathcal{L}$ are the LO and 1-loop matrix elements respectively. The LO amplitude always enters at least linearly, which ensures that the behavior of the virtual correction is often similar to the LO one. One can drastically decrease the number of evaluations of $\mathcal{L}$ by performing optimization steps using the LO matrix element $\mathcal{M}$ only. To this end, a (partially) unweighted sample of events[7] is generated at LO that follows the random distribution

$$g(\mathbf{X}) = g(x_1, x_2, \phi_m) = \frac{1}{\hat{\sigma}_{ab}^{\rm LO}} \frac{d\hat{\sigma}_{ab}^{\rm LO}}{dx_1 dx_2 d\phi_m} \,. \tag{4.72}$$

Observables $\mathcal{O}$ can be evaluated at LO using the event sample

$$\langle \mathcal{O} \rangle = \frac{1}{N} \sum_{i=1}^{N} \mathcal{O}(\mathbf{X}_i) \xrightarrow{N\to\infty} \int d\mathbf{X}\, g(\mathbf{X})\, \mathcal{O}(\mathbf{X}) \,, \tag{4.73}$$

where N is the number of events in the sample. One can then set the observable in question to the 1-loop correction, i.e.

$$\mathcal{O}(\mathbf{X}) = \left(1 + \frac{\mathcal{M}^*\mathcal{L} + \mathcal{M}^*\mathcal{L}}{|\mathcal{M}|^2}\right) \,, \tag{4.74}$$

which then yields an estimate of the virtual correction

$$\frac{1}{N} \sum_{i=1}^{N} \mathcal{O}(\mathbf{X}_i) = \frac{\sigma_{ab}^{\rm LO+V}}{\sigma_{ab}^{\rm LO}} \,. \tag{4.75}$$

The same argument holds for evaluating the virtual corrections differentially. In practical terms, each LO event is reweighted according to

$$w_{\rm LO+V} = w_{\rm LO} \left(1 + \frac{\mathcal{M}^*\mathcal{L} + \mathcal{M}^*\mathcal{L}}{|\mathcal{M}|^2}\right) \,, \tag{4.76}$$

which yields a weighted sample of events that can then be used to generate histograms of observables at 1-loop accuracy using Eq.(4.73)[8]. The phase-space points as well as the helicity and color information are provided to HELAC-1LOOP using a modified version of the LES HOUCHES FILE format [218] and written back into one after the reweighting process is completed.

[7] A partially unweighted sample will include a small number of weighted events. The reweighting method requires little to no modification in this case.

[8] The weight $w_{\rm LO}$ is 1 for fully unweighted events, but is made explicit here for the case of partially unweighted events.

4.3 Real-emission corrections

The UV divergences arising in virtual corrections are removed by renormalization. With that one still retains the soft- and collinear divergences that arise in the IR sector. Their contribution partially cancels between the virtual and real corrections, when considering IRC-safe observables. The remaining collinear singularities are then cancelled by a redefinition of parton distribution functions [219]. For analytical calculations one can explicitly perform this cancellation by introducing suitable regulators. However, automated NLO calculations typically perform the integration over the phase-space of the external particles using Monte Carlo methods. This complicates the issue, because the virtual corrections live in a m-particle phase-space, while the real-emission lives in the $m+1$ phase-space. Consequently, the use of Monte Carlo methods requires that both integrands are already finite before any integration is performed. The HELAC-NLO framework adopts subtraction scheme techniques for this purpose.

4.3.1 Subtraction schemes

Subtraction schemes were initially proposed in full generality in Ref. [219] for the massless case and later extended to the massive case in Ref. [220]. A general NLO partonic cross-section can be written as

$$\hat{\sigma}^{\rm NLO} = \int d\hat{\sigma}^{\rm NLO} = \int_m d\hat{\sigma}^{\rm LO+V} + \int_{m+1} d\hat{\sigma}^{\rm R} + \int_m d\hat{\sigma}^{\rm C}\,, \tag{4.77}$$

where $d\hat{\sigma}^{\rm LO+V}$ is the differential cross-section originating from the Born calculation and the renormalized virtual contribution, $d\hat{\sigma}^{\rm R}$ is the genuine real-emission correction and the PDF induced counterterm is made explicit by $d\hat{\sigma}^{\rm C}$. The latter arises due to collinear splitting from the initial state partons, which remain uncancelled even after summing the virtual and real contributions. The goal is now to separately cancel the singularities in the m and $(m+1)$ integrations. To achieve this, one adds and subtracts an auxiliary term $d\hat{\sigma}^{\rm A}$ that exactly reproduces the singular structure of the real-emission term, but can otherwise be chosen to be simpler. With a suitable definition one can factorize the $(m+1)$-particle integration of the auxiliary term into an m-particle and a 1-particle integration. By first performing the integration over the 1-particle phase space through either analytical or numerical means one can then include the same term in both integrations. Notably, the pole structure has to be extracted analytically even when the finite part is obtained through numerical techniques. The partonic cross section can then be written as

$$\begin{aligned}\hat{\sigma}^{\rm NLO} &= \int_{m+1}\left[d\hat{\sigma}^{\rm R} - d\hat{\sigma}^{\rm A}\right] + \int_{m+1} d\hat{\sigma}^{\rm A} + \int_m d\hat{\sigma}^{\rm LO+V} + \int_m d\hat{\sigma}^{\rm C} \\ &= \int_{m+1}\left[d\hat{\sigma}^{\rm R} - d\hat{\sigma}^{\rm A}\right] + \int_m \left[\int_1 d\hat{\sigma}^{\rm A} + d\hat{\sigma}^{\rm LO+V} + d\hat{\sigma}^{\rm C}\right]\,.\end{aligned} \tag{4.78}$$

The first term is finite by definition and can be directly integrated with Monte Carlo methods. The second one is only finite for IR-safe observables. In practice, we compute jet-observables, which are defined by a measurement function $F_J^{(m)}$ with the property that

$$F_J^{(m+1)} \to F_J^{(m)}\,, \tag{4.79}$$

whenever the kinematics of the m and $(m+1)$ particle phase-space differ by collinear and/or soft emissions only. Of course, the actual construction of the subtraction term $d\sigma^{\mathrm{A}}$ still remains. Several choices are available. HELAC-NLO implements the Catani-Seymour [219, 220] and Nagy-Soper subtraction [195, 196] schemes. For the construction of the subtraction terms one first needs to consider the singularity structure of the real-emission amplitude. The application to Monte Carlo codes further requires that the form of $d\hat{\sigma}^{\mathrm{A}}$ is independent on the observable in question. Beyond the subtraction terms $d\hat{\sigma}^{\mathrm{A}}$ one also has to supply a momentum mapping from the $(m+1)$-particle phase-space to the m-particle one or vice versa. This is needed, because a decay

$$P \to p_1 + p_2 \tag{4.80}$$

does not fulfil momentum conservation if all three particles are massless and on-shell. Thus, the remaining momentum must be absorbed into the kinematics of other particles in the process, while ensuring that observables are inclusive with respect to this variation in event kinematics.

4.3.2 Matrix elements in the soft limit

It is well-known that that the real-emission amplitude factorizes in the limit of soft emissions. The factorization of a tree-like matrix element M_{m+1} is given by eikonal currents

$$|M_{m+1}|^2 \sim {}_m\Big\langle ... \Big| (\boldsymbol{J}^{\mu,a})^\dagger \boldsymbol{J}^a_\mu \Big| ... \Big\rangle_m \,, \tag{4.81}$$

where the eikonal currents are given for a soft gluon with momentum k by

$$\boldsymbol{J}^{\mu,a} = \sum_i \boldsymbol{T}^a_i \frac{p_i^\mu}{p_i \cdot k} \,. \tag{4.82}$$

The generators of the color algebra is given by the $\boldsymbol{T}^a_{bc} = i f^{bac}$ if the i-th parton is in the adjoint representation, by $\boldsymbol{T}^a_{\alpha\beta} = t^a_{\alpha\beta}$ if it is in the fundamental representation and by $\boldsymbol{T}^a_{\alpha\beta} = -t^a_{\beta\alpha}$ in the anti-fundamental one. For polarized soft partons one finds that the eikonal approximation is modified by

$$(\boldsymbol{J}^{\mu,a})^\dagger \boldsymbol{J}^a_\mu \;\to\; -(\boldsymbol{J}^{\mu,a})^\dagger \boldsymbol{J}^{\nu,a} \epsilon_\mu(k,\lambda) \epsilon^*_\nu(k,\lambda) \,, \tag{4.83}$$

where ϵ is the polarization vector of a gluon with momentum k and polarization λ. For helicity states one can relate different polarization states by

$$\epsilon^*_\mu(k,+) = \exp(i\psi)\epsilon_\mu(k,-) \,, \tag{4.84}$$

and, using the self-conjugate property of the eikonal current, find that

$$-(\boldsymbol{J}^{\mu,a})^\dagger \boldsymbol{J}^{\nu,a} \epsilon_\mu(k,\lambda) \epsilon^*_\nu(k,\lambda) = \frac{1}{2} (\boldsymbol{J}^{\mu,a})^\dagger \boldsymbol{J}^{\nu,a} \,, \tag{4.85}$$

which implies that the soft limit is independent of the polarization vectors for helicity states [195].

4.3.3 Matrix elements in the collinear limit

Consider a pair of outgoing massless partons $\{i, j\}$ that are studied in the collinear limit. One can parametrize their momenta p_i and p_j according to

$$\begin{aligned} p_i^\mu &= z p^\mu + k_\perp^\mu - \frac{k_\perp^2}{z}\frac{n^\mu}{2p\cdot n}\,, \\ p_j^\mu &= (1-z) p^\mu - k_\perp^\mu - \frac{k_\perp^2}{1-z}\frac{n^\mu}{2p\cdot n}\,, \end{aligned} \tag{4.86}$$

where $p^2 = m_{ij}^2 = 0$, which implies that $p \neq p_i + p_j$ [9]. Furthermore, n^μ is an auxiliary light-like vector and $k_\perp$ is transverse to n and p. The collinear limit is then defined by taking $k_\perp \to \alpha k_\perp$ with $\alpha \to 0$. The squared matrix element is given by

$$\begin{aligned} &{}_{m+1}\langle ..., \{p_i, \lambda_i\}, ..., \{p_j, \lambda_j\}, ... | ..., \{p_i, \lambda_i\}, ..., \{p_j, \lambda_j\}, ... \rangle_{m+1} \\ &\xrightarrow{\alpha\to 0} \frac{1}{\alpha^2}\frac{4\pi\alpha_s}{p_i p_j}\, {}_m\Big\langle ..., \{p, \lambda'\}, ... \Big| \hat{P}_{\tilde{ij},i}(z, k_\perp, \{\lambda\}) \Big| ..., \{p, \lambda\}, ... \Big\rangle_m\,, \end{aligned} \tag{4.87}$$

where the set of helicity states is denoted by $\{\lambda\}$, $\tilde{ij}$ is the emitter parton and $\hat{P}$ are the generalized Altarelli-Parisi splitting kernels [195, 219]. The splitting kernels are given in Ref. [195] by

$$\langle \lambda' | P_{qq}(z, k_\perp, \lambda_q, \lambda_g) | \lambda \rangle = C_F \left\{ \frac{\delta_{\lambda\lambda_q}(z^2 + \delta_{\lambda_q\lambda_g}(1-z^2))}{1-z} \right\} \delta_{\lambda'\lambda}\,, \tag{4.88}$$

$$\begin{aligned} &\langle \lambda' | P_{gq}(z, k_\perp, \lambda_q, \lambda_{\bar{q}}) | \lambda \rangle = \\ &\qquad T_R \Bigg\{ \frac{\delta_{\lambda'\lambda}(1 - \delta_{\lambda_q\lambda_{\bar{q}}})}{2} \\ &\qquad\quad - \frac{(1-\delta_{\lambda_q\lambda_{\bar{q}}})(k_\perp\cdot\epsilon(p,\lambda'))^*(k_\perp\cdot\epsilon(p,\lambda)}{p_q\cdot p_{\bar{q}}} \\ &\qquad\quad + \delta_{\lambda'\lambda}\left[(1-\delta_{\lambda_q\lambda_{\bar{q}}})\left(\delta_{\lambda\lambda_q} z + \delta_{\lambda\lambda_{\bar{q}}}(1-z) - \frac{1}{2}\right)\right] \Bigg\}\,, \end{aligned} \tag{4.89}$$

and

$$\begin{aligned} &\langle \lambda' | P_{gg}(z, k_\perp, \lambda_i, \lambda_j) | \lambda \rangle = \\ &\qquad C_A \Bigg\{ \delta_{\lambda'\lambda}\left(\frac{\delta_{\lambda\lambda_i}}{1-z} + \frac{\delta_{\lambda\lambda_j}}{z} - 2(1-\delta_{\lambda_i\lambda_j})\right) \\ &\qquad\quad - 2(1-\delta_{\lambda_i\lambda_j}) z(1-z)\frac{(k_\perp\cdot\epsilon(p,\lambda'))^*(k_\perp\cdot\epsilon(p,\lambda))}{k_\perp^2} \\ &\qquad\quad + \delta_{\lambda'\lambda}[\delta_{\lambda\lambda_i} - \delta_{\lambda\lambda_j}](1-2z) \Bigg\}\,, \end{aligned} \tag{4.90}$$

where it is understood that the first index in $\hat{P}_{ij}$ indicates the emitter and the information on one of the split partons on the second index fully defines the type of splitting in QCD.

[9] This inequality necessitates the introduction of a momentum-mapping to absorb the remaining momentum.

4.3.4 Catani-Seymour subtraction

We will now focus first on the Catani-Seymour (CS) subtraction scheme [219, 220]. The auxiliary cross section can be decomposed into a sum of N dipole subtraction terms

$$d\hat{\sigma}^{\mathrm{A}}_{ab}(\{p\}_{m+1}) = \sum_{i=1}^{N} d\hat{\sigma}^{\mathrm{B}}_{ab}(\{\tilde{p}\}^{(i)}_{m}) \otimes dV_i(\{p\}_{m+1})\,, \tag{4.91}$$

where the convolution $\otimes$ assures consistency of helicities and colors, the set of unmapped momenta is $\{p\}_{m+1}$ and the set of mapped momenta for the i-th dipole term is denoted by $\{\tilde{p}\}^{(i)}_{m}$. On the other hand, the m-particle integration can be written in terms of

$$\begin{aligned}&\int_m \left[d\hat{\sigma}^{\mathrm{LO+V}}_{ab} + d\hat{\sigma}^{\mathrm{C}}_{ab} + \int_1 d\hat{\sigma}^{\mathrm{A}}_{ab}\right] \\ &= \int_m \left[d\hat{\sigma}^{\mathrm{LO+V}}_{ab} + d\hat{\sigma}^{\mathrm{LO}}_{ab} \otimes \mathbf{I}\right] + \int_0^1 dx \int_m d\hat{\sigma}^{\mathrm{LO}}_{a'b'} \otimes \mathbf{KP}_{aa',bb'}\,,\end{aligned} \tag{4.92}$$

where the first bracket contains the left-over IR singularities in the renormalized 1-loop amplitude as well as the logarithms with respect to the renormalization scale. The second term contains the explicit dependence on the factorization scale. The first bracket depends only on m-particle kinematics $\{p\}_m$, which assures that pointwise cancellation of the divergences is achieved. At NLO QCD the second term can be written as

$$\begin{aligned}&\int_0^1 dx \int_m d\hat{\sigma}^{\mathrm{LO}}_{a'b'} \otimes \mathbf{KP}_{aa',bb'} \\ &= \sum_c \int_0^1 dx \int_m d\hat{\sigma}^{\mathrm{LO}}_{cb}(xp_a, p_b, \{p\}_m) \otimes \mathbf{KP}_{ac}(x, p_a, \{p\}_m, \mu_F) \\ &+ \sum_c \int_0^1 dx \int_m d\hat{\sigma}^{\mathrm{LO}}_{ac}(p_a, xp_b, \{p\}_m) \otimes \mathbf{KP}_{bc}(x, p_b, \{p\}_m, \mu_F)\,,\end{aligned} \tag{4.93}$$

where we have used that at most one QCD splitting can take place. The explicit form of the $\mathbf{I}$ and $\mathbf{KP}$ operators as well as the dipole counterterms are given in Ref. [219, 220]. The dipoles can categorized into four topologies depending on whether the partons involved are in the initial or final state of the process:

$$\mathcal{D} = \sum_{\{i,j\}} \sum_{k\neq i,j} \mathcal{D}_{ij,k} + \sum_{\{i,j\}} \sum_{a} \mathcal{D}^{a}_{ij} + \sum_{\{a,i\}} \sum_{j\neq i} \mathcal{D}^{ai}_{j} + \sum_{\{a,i\}} \sum_{b\neq a} \mathcal{D}^{ai,b}\,, \tag{4.94}$$

where the initial state partons are labeled a, b and the final state partons by i, j and k. In Eq. (4.94), labels with two indices such as ij are understood to be the emitter - emitted pair and the lone index belongs to the spectator. Taking the example with both emitter-pair partons being in the final state of the process, the sum of the momenta of emitter and emitted partons $p_i + p_j$ will not be on-shell, when p_i and p_j are on-shell, respectively. Thus, one absorbs the additional recoil into a separate spectator parton, while being careful to preserve momentum conservation. It is desirable to do this in such a way that [128]:

- The initial state partons stay aligned with the beam-axis.
- The direction of the spectator parton remains unaffected.

The implementation is straightforward when all partons are in the final state. The mapped momenta are given by [219]

$$\tilde{p}_k = \frac{1}{1-y_{ij,k}} p_k \,, \quad \tilde{p}_{ij} = p_i + p_j - \frac{y_{ij,k}}{1-y_{ij,k}} p_k \,, \tag{4.95}$$

where the recoil parameter is given by

$$y_{ij,k} = \frac{p_i p_j}{p_i p_j + p_j p_k + p_k p_i} \,. \tag{4.96}$$

Note that this mapping ensures momentum conservation:

$$p_i + p_j + p_k = \tilde{p}_{ij} + \tilde{p}_k \,. \tag{4.97}$$

Similarly, if only the spectator is in the initial state, then one can find an expression that just rescales the initial state momentum to account for longitudinal recoil:

$$\tilde{p}_a = x_{ij,a}\, p_a \,, \quad \tilde{p}_{ij} = p_i + p_j - (1 - x_{ij,a})\, p_a \,, \tag{4.98}$$

where now the recoil parameter $x_{ij,a}$ is defined by

$$x_{ij,a} = \frac{p_i p_a + p_j p_a - p_i p_j}{(p_i + p_j) p_a} \,. \tag{4.99}$$

For initial state emitter with final state spectator one can no longer maintain both restrictions. Instead, the spectator direction is changed to compensate, while the initial state emitter stays aligned with the beam pipe:

$$\tilde{p}_{ai} = x_{ij,a} p_a \,, \quad \tilde{p}_j = p_j + p_i - (1 - x_{ij,a} p_a) \,, \tag{4.100}$$

where $x_{ij,a}$ is given by Eq. (4.99). Most problematic is the case with initial state emitter and spectator. The emitter can be kept aligned with the beam-pipe, but that would necessarily move the initial state spectator out of it. To avoid this one instead performs a Lorentz boost on all final-state particles, while keeping the spectator momentum unchanged [128, 219]. In practice, the mapping for the emitter pair is given by

$$\tilde{p}_{ai} = x_{i,ab} p_a \,, \quad x_{i,ab} = \frac{p_a p_b - p_i p_a - p_i p_b}{p_a p_b} \,, \tag{4.101}$$

and the rest of the final state momenta q are modified according to

$$\begin{aligned} \tilde{q} &= q - 2\frac{q\cdot(K+\tilde{K})}{(K+\tilde{K})^2}(K+\tilde{K}) + 2\frac{q\cdot K}{K^2}\tilde{K} \,, \\ K &= p_a + p_b - p_i \,, \qquad \tilde{K} = \tilde{p}_{ai} + p_b \,. \end{aligned} \tag{4.102}$$

It should be noted that the number of mappings is the same as if only one spectator momentum were modified. In general, the number of mappings grows as n^3, where n is the number of QCD partons. The Catani-Seymour subtraction scheme can be extended to hold on the level of helicity states as described in Ref. [195]. As previously mentioned, the soft limit stays invariant, while the collinear contribution has to be modified to reference the polarisations vectors. The implementation of the CS scheme in HELAC-DIPOLES is based on this polarized formulation. This

makes it possible to sum the helicity states of external particles using Monte Carlo techniques.

An additional optimization can be made by restricting the phase-space of the subtraction terms. Since we are only interested in soft and collinear regions, we can integrate the dipole terms over a smaller phase space, while making sure that all singular regions are still included. Because the auxiliary cross-section was introduced by adding and subtracting the subtraction term, any physical observable must be independent of the size of the phase-space in these terms [221]. In practice, this is achieved by restricting the phase space based on the recoil parameters and momenta:

$$\begin{aligned} y_{ij,k} &< \alpha_{max} \quad \text{for} \quad D_{ij,k}\,, \\ 1 - x_{ij,a} &< \alpha_{max} \quad \text{for} \quad D^{a}_{ij}\,, \\ \frac{p_i p_a}{p_i p_a + p_j p_a} &< \alpha_{max} \quad \text{for} \quad D^{ai}_{j}\,, \\ \frac{p_a p_i}{p_a p_b} &< \alpha_{max} \quad \text{for} \quad D^{ai,b}\,. \end{aligned} \tag{4.103}$$

As previously mentioned, the final result must be independent of the actual value of α_{max}, which can be chosen between $(0, 1]$. When the kinematics above exceed the value of α_{max} then the dipole is not calculated, which eliminates the need to recompute a matrix element for the respective momentum mapping. This can constitute a substantial reduction in CPU time when smaller values are chosen. In addition, one can use the independence of the exact value as a cross-check on the correct implementation of the calculation. Such checks are vital in the development process of the code, but also provide an efficient test on the selection of partonic subprocesses that contribute to a given hadronic process.

4.3.5 Nagy-Soper subtraction

The Nagy-Soper (NS) subtraction scheme [196] is an alternative scheme for the construction of dipole terms and their respective momentum mappings. The main feature is a reduction of the number of momentum mappings that have to be evaluated. Asymptotically only n^2 instead of n^3 mappings are needed. Each mapping amounts to one independent evaluation of an m-particle matrix elements, which can induce a significant CPU-time cost for processes with many QCD partons in the final state. In the context of processes with many final-state particles, this becomes strongly desirable, because the evaluation of matrix elements amounts to a large part of the computational load. This makes the Nagy-Soper scheme preferable when studying top-quark decays.

The NS formalism is constructed by defining subtraction terms for each pair of two partons with momenta p_i and p_j at a time. All soft and collinear limits of the two are included. Similarly to the CS case, a momentum mapping must be performed. Instead of selecting one spectator among the remaining partons, one performs a Lorentz boost on the entire remaining system. Following Ref. [196] we define $Q = p_a + p_b$, where p_a and p_b are the initial state momenta, and additional four-momenta

$$\begin{aligned} P_{ij} &= p_i + p_j\,, \\ K &= Q - P_{ij}\,, \end{aligned} \tag{4.104}$$

when $i \notin \{a, b\}$ as well as

$$\begin{aligned} P_{ij} &= p_i - p_j \,, \\ K &= Q - p_j \,, \end{aligned} \tag{4.105}$$

when $i \in \{a, b\}$. The momentum K constitutes a spectator that is constructed from all final-state momenta that are not part of the singularity structure itself. The mapping is then defined in the singular limit as

$$\begin{aligned} P_{ij} &\to \tilde{p}_i \,, \\ p_k &\to \tilde{p}_k \,, \quad k \neq i \,. \end{aligned} \tag{4.106}$$

The requirement that all particles, including the combined particle $\tilde{ij}$, remain on-shell, even away from the singular limit, allows for only one solution up to an additional rotation. Explicitly, given the two vectors

$$\begin{aligned} K &= Q - P_{ij} \,, \\ \tilde{K} &= \tilde{Q} - \tilde{p}_i \,, \end{aligned} \tag{4.107}$$

together with the condition $\tilde{Q} = Q$, ultimately leads to a Lorentz-transformation [196]

$$\Lambda(K, \tilde{K})^{\mu}_{\nu} = g^{\mu}_{\nu} - 2\frac{(K + \tilde{K})^{\mu}(K + \tilde{K})_{\nu}}{(K + \tilde{K})^2} + 2\frac{K^{\mu}\tilde{K}_{\nu}}{K^2} \,. \tag{4.108}$$

An additional rotation around the vector $\tilde{K}$ would similarly lead to a valid solution. For the case with an initial state parton splitting one can no longer enforce $Q = \tilde{Q}$ and instead requires that the initial states don't obtain any transverse momentum. This yields the same form for the Lorentz transformation, but the vectors are instead given by

$$\begin{aligned} K &= p_a + p_b - p_j \,, \\ \tilde{K} &= \tilde{p}_a + \tilde{p}_b = \tilde{p}_a + p_b \,, \end{aligned} \tag{4.109}$$

where the momentum of the splitting parton $\tilde{p}_a$ is fixed through the requirements $K^2 = \tilde{K}^2$ and $\tilde{p}_a = x p_a = \frac{K^2}{Q^2} p_a$. With this mapping in place one can schematically write the subtraction term as

$$d\sigma^{\mathrm{A}} = \sum_{i,j,k}^{m+1} d\sigma^{\mathrm{B}}(\{\tilde{p}\}_m^{(ij)}) \otimes D^{(ijk)}(\{\tilde{p}\}_m^{(ij)}, \{p\}_{m+1}) \,, \tag{4.110}$$

where it has been made explicit that the mapped momenta are derived for each pair of $\{i, j\}$ and do not depend on k. The exact form of the subtraction kernels is derived in Ref. [196] and is further discussed for the QED case in [209]. The integration over the unresolved phase space is performed using a semi-numerical method. Using the general pole structure of the soft and collinear limits one can extract the $1/\varepsilon$ and $1/\varepsilon^2$ poles on the integrand level and then solve the remaining integrals numerically. This forgoes the need for a difficult analytical solution induced by the more complicated momentum mapping.

HELAC-DIPOLES also implements a restriction parameter on the phase space of the dipole subtraction terms similarly to the Catani-Seymour case, which can again be used for internal consistency checks. In addition, the presence of two subtraction schemes in the program allows for consistency checks using both. Any physical amplitude does not, of course, depend on the choice of the subtraction scheme. Therefore the real-emission contribution must be the same regardless of which scheme was used.

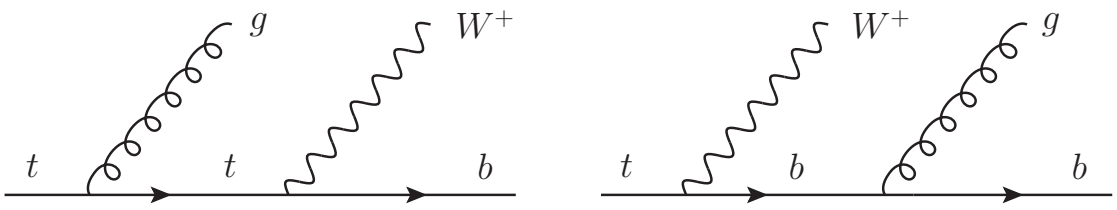


Figure 4.4: *The two real-emission contributions to the NLO QCD top-quark decay into a W boson and a bottom quark. The gluon can be emitted either from the top or the bottom line.*

4.3.6 Subtraction for top-quark decays in the NWA

The inclusion of top-quark decays in the NWA introduces three divergent contributions in the real-emission contribution. Consider an on-shell top-quark decaying to a W-boson and a massless bottom quark:

$$t \to W^+ bg \tag{4.111}$$

Two real-emission diagrams contribute to the real-emission matrix element. The diagrams for the decay are shown in Figure 4.4. A soft gluon emission from the top-quark line generates a soft divergence. No potential collinear divergence appears here due to the mass of the top-quark. On the other hand, the gluon emission from the massless b-line generates both a soft and a collinear singularity, which is fully in line with the previously discussed case of final-state IR singularities. It is possible to construct a subtraction term that cancels both divergences as described in Ref. [222], which we will closely follow in this section. The momentum mapping has to be performed with caution. The W-boson will enter a subsequent decay, which introduces a Breit-Wigner propagator to the matrix element. To avoid the large variation in the Breit-Wigner distribution, one must ensure that the mapped W-boson momentum fulfills $\tilde{p}_W^2 = p_W^2$. In addition, both the top and bottom quarks are required to be on-shell. The momentum conservation constraints $p_t = \tilde{p}_W + \tilde{p}_b$ further restrict the form of the mapping. Generally, p_W and $\tilde{p}_W$ are related by a Lorentz transformation. The mapped W-boson momentum can be expressed in terms of the unmapped momenta through

$$\tilde{p}_W = \alpha \left(p_W - \frac{p_t \cdot p_W}{p_t^2} p_t \right) + \beta \, p_t \,, \tag{4.112}$$

where

$$\alpha = \frac{p_t^2 - p_W^2}{2\sqrt{(p_t \cdot p_W)^2 - p_W^2 p_t^2}} \,, \quad \beta = \frac{p_t^2 + p_W^2}{2p_t^2} \,. \tag{4.113}$$

The exact form of the subtraction term is given Ref. [222]. HELAC-DIPOLES implements the subtraction of the singularity of intermediate top quarks in the NWA following the procedure described here.

4.4 Reweighting event files

The estimation of theoretical uncertainties requires many re-computations of the simulations. For example, the internal uncertainty for PDF sets is usually obtained

by evaluating the observable with $\mathcal{O}(100)$ PDF eigenvectors, while the uncertainties related to the fixed-order expansion in the coupling constant α_s is typically performed by evaluating each observable at 7 points of the renormalization and factorization scale and computing the envelope of the different predictions. Furthermore, the PDF4LHC working group recommends the use of at least 3 different PDF sets for SM predictions [223]. Repeating the computation for many scale and PDF settings is overwhelmingly expensive. It is more economical to instead store a (partially) unweighted event sample that can be reweighted for different setups. To that end, HELAC-NLO stores events in modified LES HOUCHES EVENT FILES [218] alongside additional information such as the momentum fractions $x_{1,2}$ and the explicit dependence on μ_R and μ_F. The relevant information depends on the component in question. We will briefly discuss the information needed for the LO contribution, virtual corrections, subtracted real, $\mathcal{I}$-operator and $\mathcal{KP}$-operator. The actual reweighting logic is implemented in the in-house HEPLOT code [200].

Naturally, the reweighting formalism at LO is the most straightforward one, but all the findings here will propagate to the NLO components. The LO weight has no explicit dependence on the renormalization scale other than through the running of $\alpha_s(\mu_R)$. In order to reweight to a different scale, one must store the value of μ_R used during the generation of the event and the powers of α_s in the squared matrix element. The new weights can then be obtained from the stored weight via simple rescaling by $\frac{\alpha_s(\mu_R)}{\alpha_s(\mu_R^0)}$, where μ_R is the desired scale and μ_R^0 is the one used in the event generation. The running of α_s is provided through the LHAPDF interface [224] and depends on the PDF set in question. Thus, the LHAPDF ID of the respective PDF set used during event generation has to be stored as well. Combined with the momentum fractions x_1 and x_2 as well as the factorization scale μ_F, one can then also easily evaluate the contribution from a new PDF set and rescale the weight accordingly. LHC calculations often neglect quark masses, which results in many sub-processes sharing the same matrix elements. In the HELAC-NLO framework this is exploited by summing the PDF functions of these before folding with the matrix elements. This heavily reduces the number of sub-processes that need to be computed. The extensions to the reweighting procedure are straightforward. The labels of the included partons in the sum are stored inside the event files, which makes the extension of the reweighting procedure trivial. One additional detail arises due to the partial unweighting approach employed throughout all HELAC codes. On top of the regular cross section value stored in each LHEF we also store a normalization factor that accounts for possibility of weighted events, which allows us to perform a partial unweighting. With this setup most events will follow the traditional unweighting approach, but a small number of weights exceeding the maximum weight can be stored as weighted events, which allows for a less precise determination of the latter to be used.

All the considerations from the LO weights carries over to the virtual corrections. Powers of α_s and the PDF information is used to rescale to different PDFs. Unlike the LO case, the 1-loop weight has an explicit dependence on μ_R. It can be parametrized in the following way. Factoring out the global powers of α_s, the finite part of the 1-loop weight can be decomposed as

$$w_{1\mathrm{L}} = c_{\mathrm{L}0} + c_{\mathrm{L}1}\log(\mu_R^2) + c_{\mathrm{L}2}\log^2(\mu_R^2)\,, \tag{4.114}$$

where the coefficients $c_{\mathrm{L}0}$, $c_{\mathrm{L}1}$ and $c_{\mathrm{L}2}$ are constants with respect to μ_R. In principle,

one can extract the constant coefficients by reevaluating each phase-space point at three scales $\mu_R^{(i)}$ and solving the resulting system of equations. The original implementation in HELAC-1LOOP indeed implements this method. However, throughout this thesis this behavior was replaced. Instead one can make use of the relation between the logarithmic structure and the pole structure to directly get the coefficients with no additional computational strain. Consider a generic unrenormalized amplitude $\mathcal{L}^{\mathrm{B}}$ in $D = 4 - 2\varepsilon$ dimensions. The integral over the loop numerator $\bar{N}(\bar{l})$ can be expanded to

$$\begin{aligned}
\mathcal{L}^{\mathrm{B}} =& \mu_R^{2\varepsilon} \int d^d l\, \bar{N}(\bar{l}) \\
=& \left[1 + \varepsilon \log(\mu_R^2) + \frac{1}{2}\varepsilon^2 \log^2(\mu_R^2) + \mathcal{O}(\varepsilon^3)\right] \left[\frac{1}{\varepsilon^2} a^{(2)} + \frac{1}{\varepsilon} a^{(1)} + a^{(0)}\right] \\
=& \frac{1}{\varepsilon^2} a^{(2)} + \frac{1}{\varepsilon}\left(a^{(1)} + a^{(2)} \log(\mu_R^2)\right) \\
&+ \left(a^{(0)} + \log(\mu_R^2)\, a^{(1)} + \frac{1}{2}\log^2(\mu_R^2)\, a^{(2)}\right) + \mathcal{O}(\varepsilon) \\
=& \left[\frac{1}{\varepsilon^2} \hat{w}_{1\mathrm{L}}^{(2)} + \frac{1}{\varepsilon} \hat{w}_{1\mathrm{L}}^{(1)} + \hat{w}_{1\mathrm{L}}^{(0)}\right],
\end{aligned} \tag{4.115}$$

where the weights $\hat{w}_{1\mathrm{L}}^{(k)}$ do not include the UV counterterms and indicate the component proportional to ε^{-k} in the expansion of the regulator. Note also that the weight of the virtual corrections is a linear function of the loop amplitude. Therefore the method can be applied on both levels with no further modification. The structure of HELAC-1LOOP naturally provides the exact pole structure, because the scalar integrals returns it alongside the finite remainder. Using Eq. (4.115) then provides a straightforward recipe to extract the logarithmic coefficients. One first extracts the squared logarithmic components, then uses that result to obtain the linear one and finally obtains the constant term via

$$\begin{aligned}
c_{\mathrm{L}2} &= \frac{1}{2} a^{(2)} = \frac{\hat{w}_{1\mathrm{L}}^{(2)}}{2}, \\
c_{\mathrm{L}1} &= a^{(1)} = \hat{w}_{1\mathrm{L}}^{(1)} - \log(\mu_R^2)\, \hat{w}_{1\mathrm{L}}^{(2)}, \\
c_{\mathrm{L}0} &= \hat{w}_{1\mathrm{L}}^{(0)} - \log(\mu_R^2)\, c_{\mathrm{L}1} - \log^2(\mu_R^2)\, c_{\mathrm{L}2}.
\end{aligned} \tag{4.116}$$

The HELAC-1LOOP program stores the weights of the counterterms summed together with the 1-loop amplitude. In general, the logarithmic structure of the counterterms depends on the renormalization conditions of the chosen renormalization scheme. For generality, we fit this part separately by evaluating three scales explicitly. The counterterms enter only as a Born-like contribution, which makes the CPU time needed for their evaluation negligible compared to the full one-loop amplitude.

To summarize, we store the LO and 1-loop weights alongside their respective coupling powers. For the 1-loop case we also store the coefficients in logarithmic expansion as previously described. The information on the PDFs is identical to the LO case and requires no modification. Finally, the normalization factors are also included, but correspond to the LO event sample before the reweighting. The 1-loop events are combined with this factor as-is to recover the correct 1-loop cross section.

The subtracted-real contribution contains the PDF information just like the LO case. The subtracted-real events store the information of the real-emission contributions alongside all corresponding dipole subtraction terms. Rather than storing

the mapped event kinematics for each dipole term, only the $(m+1)$-kinematics are stored and the momentum mapping is recomputed in HEPLOT during the reweighting process. To that end, the subtraction scheme, the number of subtraction terms and the pair (triplet) of indices characterizing the dipole term in Eq. (4.94) are provided for the NS (CS) subtraction scheme. Modern Monte Carlo simulations often employ dynamical scale choices for μ_R and μ_F to improve the perturbative stability of differential distributions. Each subtraction term comes with its own kinematics, which can result in different scales for each term, and by extension, different values of α_s. Consequently, we store scales and weights for each subtraction term separately to allow for independent reweighting of their respective contributions. Combined with the information on the momentum mapping we can reweight all weights separately and finally recombine them to obtain the reweighted total weight. Lastly, information on the jet algorithm itself and the outcome of the jet algorithm for specific events is used to identify the jet observables that constitute IRC-safe observables in the context of subtraction schemes.

The $\mathcal{I}$-operator contains the subtraction term that is needed to cancel the IR divergences in the virtual corrections. Finite contributions dependent on the logarithms of μ_R are contained within and have to be provided explicitly from the integrated dipoles. In practice, we find the same structure as in the 1-loop case. Three coefficients have to be provided that contain constant, logarithmic and squared logarithmic contributions. The form of the coefficients is extracted analytically and implemented directly in HELAC-DIPOLES. However, one could also implement a fitting procedure instead. The information of the PDF set and the α_s powers is identical to the LO case otherwise.

Lastly, the $\mathcal{KP}$-operator information is more complicated due to the non-diagonal elements in flavor space. The logarithmic structure is decomposed in a similar way to the 1-loop case. We write

$$\begin{aligned} w(\mu_F) = &\left(L_0^{A1_g} + L_1^{A1_g}\,\log(\mu_F^2)\right) F_{A1_g} + \left(L_0^{A1_q} + L_1^{A1_q}\,\log(\mu_F^2)\right) F_{A1_q} \\ &+ \left(L_0^{B1_g} + L_1^{B1_g}\,\log(\mu_F^2)\right) F_{B1_g} + \left(L_0^{B1_q} + L_1^{B1_q}\,\log(\mu_F^2)\right) F_{B1_q} \\ &+ \left(L_0^{A2_g} + L_1^{A2_g}\,\log(\mu_F^2)\right) F_{A2_g} + \left(L_0^{A2_q} + L_1^{A2_q}\,\log(\mu_F^2)\right) F_{A2_q} \\ &+ \left(L_0^{B2_g} + L_1^{B2_g}\,\log(\mu_F^2)\right) F_{B2_g} + \left(L_0^{B2_q} + L_1^{B2_q}\,\log(\mu_F^2)\right) F_{B2_q}\,, \end{aligned} \tag{4.117}$$

where the coefficients F contain the dependence on the parton-density functions. In particular,

$$\begin{aligned} F_{A1} &= \sum_{a'} f_{a'}(x_a/x'_a,\mu_F)/x'_a\; f_b(x_b,\mu_F)\,, \\ F_{B1} &= \sum_{a'} f_{a'}(x_a,\mu_F) f_b(x_b,\mu_F)\,, \end{aligned} \tag{4.118}$$

where the sum over a' runs only over the gluon for the the gluon coefficients and over the quarks included in the proton definition for the quark case. The functions F_{A2_g}, F_{A2_q}, F_{B2_g} and F_{B2_q} are obtained by replacing $a \leftrightarrow b$ and $a' \leftrightarrow b'$. The PDF summation is implemented by replacing F with $\bar{F}$ [209]

$$\bar{F} = \sum_{\{a,b\}} F(a,b)\,, \tag{4.119}$$

where the sum runs over all pairs of included initial-state partons in the PDF summation. The content of the sums in Eq. (4.118) is stored. The event files must include the full information on the PDF summation pairs as well as the coefficients L_0 and L_1. Furthermore, the momentum fractions x_1' and x_2' have to provided on top of the LO information. We highlight that the event information is much larger for the KP-operator compared to the other contributions. We will see that the EFT-expansion of this logic will be particularly expensive in terms of storage space as a result. However, this approach still constitutes an improvement over naively storing similar information for each partonic channel.

Chapter 5

HELAC-SMEFT

The main technical aspect of the project presented in this thesis consists of expanding the previously described HELAC-NLO program for applications to SMEFT and potentially other effective field theories. In the process large aspects of the programs have to be modified, replaced or extended. Limitations of the original constructions are identified and resolved. The generalizations involved can lead to reduced performance in practice. Several optimizations of both technical and conceptual nature have been implemented to offset such regressions. Furthermore, the structure of the codes is extended to support EFT calculations in various ways. A new interface code was developed to automatically construct vertex and current functions, which contain the model information and supersede the previous hard-coded SM implementation. Lastly, all reweighting facilities are extended to support existing functionality in the more complicated SMEFT application. In this chapter, we outline all the modifications that have been made, and touch on many conceptual and technical aspects of the new framework. The new software package is called HELAC-SMEFT and is developed with the following goals in mind. The program

- shall be flexible in the operators included in the Lagrangian,
- must support calculations with unstable particles treated either in the NWA or off-shell through the use of the complex-mass scheme,
- should consistently include operators introduced by the effective field theory both in the production and decay stage of the process,
- provide results organized by Wilson coefficients in a single run, i.e. provide results for the SM as well as linear, quadratic and interference SMEFT contributions,
- allow for a fully dynamic, per-event definition of the scale μ_{EFT} at which the operators are defined,
- store results in a suitable format, so that all events can fully be reweighted to different choices of μ_{EFT} and Λ and the running of the Wilson coefficients C_i can be retroactively inserted,
- and preserve the existing SM functionality.

In broad strokes the adjustments needed in HELAC-NLO involve adding support for vertices with more than four particles, updating the book-keeping logic to support the simultaneous calculation of several EFT contributions, and replace all SM-specific optimization strategies and assumptions with robust model-agnostic implementations. Furthermore, the modified LHE files, which are used to output the results, have to be extended to support the extended reweighting technology. Lastly, the Feynman rules of the model have to be implemented, which includes the LO vertices, as well as UV renormalization counterterms and R2 terms for NLO calculations. To lessen the impact of the increased complexity on the CPU resource consumption, we further implement additional optimizations. Starting from the implementation of the vertices induced by the effective operators, we note that a manual implementation quickly becomes very laborious due to the huge number of UV counterterms and R2 terms. Furthermore, the large scale of the development means that additional restrictions and conventions might not be known a priori, which could then instigate further rewriting of existing work as development continues. Consequently, an automated approach is needed to safely and efficiently carry out this task. To give a perspective on the complexity involved, we provide the number of vertices generated in the SM and for the five SMEFT operators $\mathcal{O}_{t\phi}$, $\mathcal{O}_{\phi G}$, $\mathcal{O}_{tG}$, $\mathcal{O}_{tW}$ and $\mathcal{O}_{tB}$ in Table 5.1. The vertices are organized in terms of the corresponding Wilson coefficients, i.e. C_{tB} was replaced with C_{tZ} as described in Chapter 3. The numbers are given based on the SMEFT@NLO UFO model [1]. The number of vertices in the SM sums to 324, while even just 5 operators increase the overall count to 1610. In principle, one can reduce the number by identifying common structures among the vertices. However, this comes at the cost of additional classification work, which would likely require an automated approach as well. One should also note that the HELAC algorithm requires all unique permutations of input currents in a given vertex separately, i.e. $\bar{u}u \to g$ and $ug \to u$ require a separate set of functions. In the context of EFT, the number of permutations drastically increases due to the increased number of external legs in vertices. As a result, the 1610 vertices outlined in this example actually decompose in many thousands of routines that need to be prepared.

Multiple modern Monte Carlo programs, such as MADGRAPH5_AMC@NLO [100], GoSam-3.0 [106], have adopted an interface to the UNIVERSAL FEYNMAN OUTPUT (UFO) format [101, 102]. Indeed, a NLO QCD implementation is available in the UFO format under the name SMEFT@NLO [105], albeit with some limitations. We choose to follow the same approach. The UFO format provides all information derived from the Lagrangian in terms of Feynman rules, particle content and coupling definitions. The Feynman rules are decomposed into several terms in color and Lorentz space and the user is expected to perform their recombination. Furthermore, the Monte Carlo program must perform all the manipulations needed to translate to their respective conventions. Finally, the resulting rules must be converted to source code and inserted into the respective code. The building blocks for this interface and the general workflow of the HELAC-SMEFT program is presented in Figure 5.1. Newly developed and extended codes are highlighted in blue. The

[1] Vertices with shared particle content have been summed, and vanishing vertices removed, for the purposes of counting. However, a fully independent implementation would require a smaller number of counterterm vertices, because the SMEFT@NLO model does not fully simplify all internal structures.

	SM	$C_{t\phi}$	$C_{\phi G}$	C_{tG}	C_{tW}	C_{tZ}	Sum
LO	155	9	12	14	66	14	270
CT	61	12	128	402	132	28	763
R2	108	10	238	221	0	0	577
Sum	324	31	378	637	198	42	1610

Table 5.1: *Number of vertices at LO, counterterms (CT) and R2 terms introduced by the SM and the five operators $\mathcal{O}_{t\phi}$, $\mathcal{O}_{\phi G}$, $\mathcal{O}_{tG}$, $\mathcal{O}_{tW}$ and $\mathcal{O}_{tB}$. The vertices induced by $\mathcal{O}_{tB}$ are split between the C_{tW} and C_{tZ} contributions via the replacement $C_{tZ} = -\sin(\theta_W)\, C_{tB} + \cos(\theta_W)\, C_{tW}$. Numbers are evaluated based on the SMEFT@NLO UFO model.*

programming language hosting the modules is also provided at the top of each block. This chapter describes the modules developed to facilitate this interface, the updates needed in the core HELAC-NLO and HEPLOT programs, as well as the multitude of optimization strategies that were implemented as part of this thesis. An overview of the framework is shown in Figure 5.1. The left-most block houses the interface from the UFO model to the HELAC-NLO. The interface is written in PYTHON, but offloads core aspects of the color-decomposition to FORM [225]. The interface block generates the FORTRAN routines and auxiliary files needed to embed the vertices and currents into the HELAC algorithm. The block is split into two main modules:

- EXPLICIT CURRENT FROM UFO (ECRU) performs analytical modifications needed to match the internal conventions in HELAC. In particular, the color-flow decomposition is implemented by dynamically generating and executing FORM code, while the construction of currents in the Lightcone representation is implemented using the tensor capabilities of SYMPY.

- HMODELGENERATOR takes the output of ECRU and generates FORTRAN90 routines for the vertex and current functions as well as additional files to embed the routines into HELAC-NLO framework.

This structure isolates the analytical conventions into ECRU module, while the details of the implementation in HELAC-NLO are isolated into HMODELGENERATOR. This split allows for easier reuse in future revisions. For example, targeting a fully new implementation of HELAC could be accomplished by simply replacing the code generation part on its own. The middle block contains the updated versions of HELAC-NLO, consisting of HELAC-1LOOP and HELAC-DIPOLES, which are compiled together with the generated routines and compute event files for the cross sections at NLO in QCD. As in previous versions, the HELAC-NLO program is implemented in FORTRAN to benefit from the strong numerical performance of the language. Finally, the event files are read by the in-house C++ program HEPLOT, which utilizes ROOT [226] to convert them into ROOT NTUPLE FILES [227]. The NTUPLE files can then be read to generate histograms [2], reweight to different

[2] HELAC-NLO also provides its own built-in histogram routine that can be utilized at runtime.

scales and PDFs and apply additional cuts [3]. An in-house PYTHON module, which utilizes the tensor manipulation capabilities of SYMPY, has also been interfaced with the framework to allow for independent checks on the UFO model in question. Feynman rules are generated based on a hand implemented Lagrangian and then provided to the ECRU module in the same way as the UFO model. This structure allows for easier cross-checks of the UFO model and hybrid setups, where specific counterterms or R2 terms can be supplied separately to complement existing UFO models. We note that the current implementation is limited to translation of vertex functions. The translation of the relations between derived model constants, such as the electroweak mixing angle, is not fully automated and the particle content is currently fixed to SM particles.

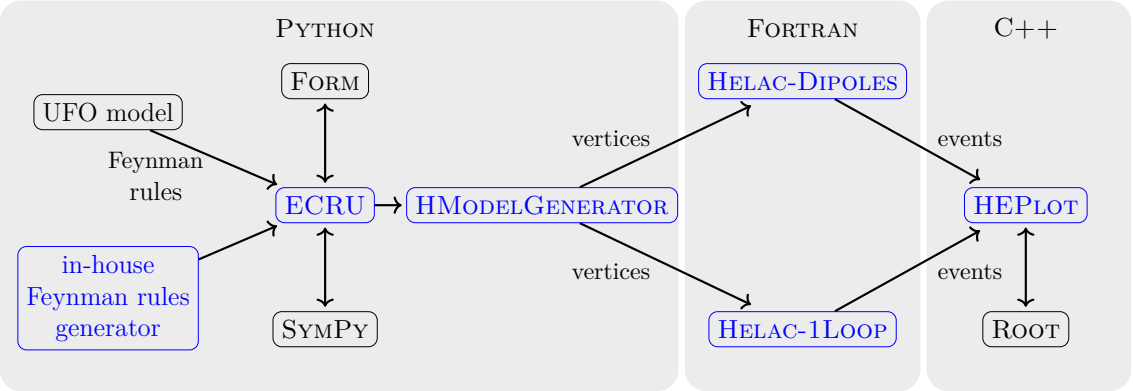


Figure 5.1: *Schematic representation of the HELAC-SMEFT framework and its workflow. Newly developed modules and updated parts of the HELAC-NLO program are highlighted in blue. The underlying programming languages are given at the top of each block. Adapted from Ref. [3].*

We first discuss additional details of the UFO interface block. Then we will detail the various modifications needed inside the HELAC algorithm as well as the additional changes implemented for HELAC-1LOOP and HELAC-DIPOLES. Lastly, we describe the extension of the LHEF format to enable the reweighting of events in the context of SMEFT.

5.1 UFO interface

As previously described, the UFO interface for HELAC-NLO is split into two distinct modules, ECRU and HMODELGENERATOR. A `user.input` allows the user to set the directory of the UFO model and provide a path to the FORM executable. The option to generate either the LO vertices or the counterterm vertices (understood to be both UV counterterms and R2 terms) is given. A filter on included couplings is provided, which allows for the selection of specific Wilson coefficients. Vector currents are natively given in the Lightcone representation in HELAC, but the option to generate them in Lorentz representation is also available for the purposes of debugging. HELAC expects all currents in two formats. One format includes the propagator of the output current within the analytical expression, while the other

[3] Cuts applied this way must always be more stringent than the ones used during the original Monte Carlo generation.

only contains the vertex structure. The former is used for all currents other than the endpoint of the recursion, where the latter must be used instead. ECRU offers the option to perform this multiplication either for the full propagator or for the numerator of the propagator only. HELAC expects the numerator only, which unifies the treatment of Born amplitudes and 1-loop numerators in the recursion. Lastly, filters are provided to discard Goldstone bosons [4] and to specify the ghosts that should be included. The code is launched using the python script `python run_ECRU.py`. The runscript can be modified to specify the number of available cores to be used by the code. Furthermore, some options can be set to generate Feynman rules only, generate the full vertex and current files in an intermediate XML-like format and to optionally print results to the screen rather than to files. Lastly, one can restrict the search to specific vertices. The code searches the model for all vertices compatible with the various restrictions and then dispatches the computation of the vertices to the number of specified jobs. The duration of the runs strongly depends on the number of vertices and the complexity of the underlying Lorentz structures. In particular, structures with products of many gamma matrices can be time-consuming to expand in SYMPY. In practice, one will usually select a specific set of operators, which are interesting in the processes to be studied and then generate the model data only once and for all. The UFO format provides the Lagrangian in the form of Feynman rules, which are decomposed into a set of not strictly unique Lorentz structures, color structures and coupling constants. The Feynman rule for a specific vertex is obtained by adding up potentially multiple contributions encoded in this format. In the first step, ECRU collects the full Feynman rules for each configuration of external legs and split them according to their perturbative coupling content. These are then divided back into their respective color structure, which can then be converted to the color-flow decomposition basis. The Lorentz-structure is similarly processed and simplified. Finally, a unique four-component current is prepared for each possible color-flow configuration allowed by the color structure of the vertex.

5.1.1 Color decomposition

The color-flow decomposition was previously described in Section 4.1.2. We will now discuss the algorithm used to automatically obtain the color-flow configurations of any vertex. Starting from the Feynman rule, one isolates the color-structures and identifies all free indices in the adjoint representation a_i. Then one contracts each index a_i with a rescaled generator matrix $\tau^a_{ij} = \sqrt{2}\, t^a_{ij}$, where t^a are the generators of the $SU(3)$ group. In the fundamental representation they are given by the Gell-Mann matrices via $t^a = \lambda^a/2$. Note that the strategy employed here works fully algebraically and never makes reference to any explicit matrix representation. Finally, one recursively applies the Fierz identity

$$\tau^a_{ij}\,\tau^b_{kl} = \delta_{il}\,\delta_{jk} - \frac{1}{3}\,\delta_{ij}\,\delta_{kl} \tag{5.1}$$

[4] Their inclusion can be discarded for NLO QCD calculations with the Higgs sector treated in the unitary gauge. However, higher order calculations in the EW coupling constants, cannot use the unitary gauge and thus have to include them.

to convert the generators into Kronecker deltas. Tensors in the adjoint space have to first be replaced by corresponding generators using

$$\begin{aligned} f^{abc} &= -\frac{i}{T_F}\left(t^a_{ij}\, t^b_{jk}\, t^c_{ki} - t^c_{ij}\, t^b_{jk}\, t^a_{ki}\right) \\ d^{abc} &= +\frac{1}{T_F}\left(t^a_{ij}\, t^b_{jk}\, t^c_{ki} + t^a_{ij}\, t^c_{jk}\, t^b_{ki}\right) . \end{aligned} \tag{5.2}$$

Finally, we obtain an expression for the Feynman rule of the vertex, which is proportional to deltas in color-space. To understand the procedure, first consider that any internal gluon line inside an amplitude is always connected to a gluon propagator that goes as δ^{ab} in the adjoint color-space. We can convert it to the fundamental representation via $\delta^{ab} = tr(\tau^a\tau^b)$ and then distribute the two generators to the connected vertices. For external lines, the algorithm automatically generates the corresponding projector that is needed to extract the physical gluon states when summing two fundamental indices $i, j \in (1, 2, 3)$. We note that the HELAC algorithm is organized such that amplitudes are categorized according to the structure of color-connections. The practical implementation of the decomposition is implemented in FORM. Given a color-structure for a specific vertex, we generate a FORM program using PYTHON, which performs the color-decomposition. The result is reinserted into the PYTHON framework to be combined with the Lorentz-structure of the vertex. To cross-check the implementation a separate fully SYMPY based module has been developed, which exploits the built-in tensor capabilities of the open-source module. In practice, we find massive performance gains in the FORM implementation.

Example: The four-gluon vertex in the SM

To illustrate the algorithm we consider the four-gluon vertex in the SM. The Feynman rule of the vertex can be written as

$$F^{\mu\nu\rho\sigma;abcd} = -ig_s^2\Big[\, f^{eab}f^{ecd}L_1^{\mu\nu\rho\sigma} + f^{eac}f^{edb}L_2^{\mu\nu\rho\sigma} + f^{ead}f^{ebc}L_3^{\mu\nu\rho\sigma}\,\Big], \tag{5.3}$$

where $L_i^{\mu\nu\rho\sigma}$ are the tensor structures of the vertex. Contracting the generators we obtain six unique products of Kronecker deltas [3]

$$\begin{aligned} F^{\mu\nu\rho\sigma;abcd}\,\tau^a_{i_1,j_1}\,\tau^b_{i_2,j_2}\,\tau^c_{i_3,j_3}\,\tau^d_{i_4,j_4} &= \delta_{i_1j_2}\delta_{i_2j_4}\delta_{i_3j_1}\delta_{i_4j_3}T_1^{\mu\nu\rho\sigma} \\ &+ \delta_{i_1j_3}\delta_{i_2j_1}\delta_{i_3j_4}\delta_{i_4j_2}T_2^{\mu\nu\rho\sigma} \\ &+ \delta_{i_1j_2}\delta_{i_2j_3}\delta_{i_3j_4}\delta_{i_4j_1}T_3^{\mu\nu\rho\sigma} \\ &+ \delta_{i_1j_4}\delta_{i_2j_1}\delta_{i_3j_2}\delta_{i_4j_3}T_4^{\mu\nu\rho\sigma} \\ &+ \delta_{i_1j_3}\delta_{i_2j_4}\delta_{i_3j_2}\delta_{i_4j_1}T_5^{\mu\nu\rho\sigma} \\ &+ \delta_{i_1j_4}\delta_{i_2j_3}\delta_{i_3j_1}\delta_{i_4j_2}T_6^{\mu\nu\rho\sigma}\,, \end{aligned} \tag{5.4}$$

where the tensors $T_i^{\mu\nu\rho\sigma}$ are defined by

$$\begin{aligned} T_1^{\mu\nu\rho\sigma} = T_2^{\mu\nu\rho\sigma} &\equiv -ig_s^2\frac{L_1^{\mu\nu\rho\sigma} - L_2^{\mu\nu\rho\sigma}}{2} \\ T_3^{\mu\nu\rho\sigma} = T_4^{\mu\nu\rho\sigma} &\equiv -ig_s^2\frac{L_3^{\mu\nu\rho\sigma} + L_1^{\mu\nu\rho\sigma}}{2} \\ T_5^{\mu\nu\rho\sigma} = T_6^{\mu\nu\rho\sigma} &\equiv -ig_s^2\frac{L_2^{\mu\nu\rho\sigma} - L_3^{\mu\nu\rho\sigma}}{2}\,. \end{aligned} \tag{5.5}$$

Eq. (5.4) can be understood as the six ways that colors can be connected inside the recursion through this vertex. In practice, only one term will apply to a given color-ordered amplitude. The skeleton generation phase of the HELAC algorithm tests any external configuration of color-flow configurations against the color-structure of the vertices and selects applicable ones if present. The subroutines that encode the current structure of the vertex implement each tensor T_i separately so that one can directly apply the selected ones without any further calculation. To be precise, for each unique color-flow i we prepare a vertex function that combines the input currents J_i

$$V_i^\mu(J_1, J_2, J_3) \equiv T_i^{\mu\nu\rho\sigma} J_1^\nu J_2^\rho J_3^\sigma \,, \tag{5.6}$$

where i directly labels the color-flow configuration. This is in contrast to the original implementation in HELAC-NLO that partially tried to label configurations based on whether they connect to the $U(1)$ gluon or not. The motivation for this will become clear shortly. However, it is clear that this approach does not work generally and thus many exceptions were present. In the updated framework, we have replaced the logic with naive labelling. However, one additional complication arises in the implementation of the $U(1)$ gluon. From the perspective of the Lagrangian it is clear that each $U(1)$ gluon propagator should carry a factor of $\frac{1}{N_C}$, which is generated by the propagator. The decomposition algorithm presented here will generate this factor directly and bind it to the currents. However, for internal propagators both connected vertices will contribute this factor, which is then double counted. On the other hand, vertices connecting to external states do not suffer from any double-counting. One observes that one cannot arrange the factor in the currents so that the overall factor is correct in both cases. We correct this mismatch during the Dyson-Schwinger recursion instead. To that end, the updated HELAC-NLO stores an additional integer for each generated current that counts how many internal $U(1)$ gluon lines are attached to it. The double counting is then corrected at runtime by replacing the output of the vertex function V with

$$V \to (-i\sqrt{N_C})^{N_{U(1)}^{\text{int}}} V \,, \tag{5.7}$$

where $N_{U(1)}^{\text{int}}$ is the number of connected $U(1)$ gluons that do not connect to any external state.

5.1.2 Lorentz structure

In HELAC all currents are treated as four-dimensional objects, where the four components correspond to the spinors for (anti-)fermions and to the four vector components for vector currents. Scalars are padded with zeros so that they fit the same structure. The tensor structures $T_i^{\mu\nu\rho\sigma}$ in the previous example were given in Lorentz representation in the previous section, but must be converted into the Lightcone representation to match the internal conventions of the HELAC-NLO framework. The procedure is straightforward, but some care must be taken. The final current functions should be functions of the four-momenta and input vector currents in Lightcone representation, but also must themselves be in Lightcone coordinates. To illustrate a current function V^μ in Lorentz representation might be a function of four-momenta p_i^μ and vector currents J^μ. The dependence on the input arguments must be replaced with the expressions for p^A and J^A. Then finally the whole expression

V^μ must be converted into the Lightcone representation to obtain $V^A(\{J^A\},\{p^A\})$, where $\{J^A\}$ and $\{p^A\}$ are the sets of input currents and momenta respectively. Finally, the numerator of the propagator, which includes the spinor structure $(\not{p}+m)$ and the Lorentz-tensor $g^{\mu\nu}$ for vectors, is also included in the current definition. The ECRU code automatically handles all these conversions after collecting the relevant structures from the Feynman rules. The replacements of the input components is straightforward. However, the conversion of the current V^μ is only performed after all other simplifications have been performed. The gamma matrices are replaced with the explicit representation in the Weyl basis. We also automatically collect left- and right-handed projectors to simplify expressions as much as possible. While SymPy provides many powerful tools for handling tensor structures, including native support for different types of indices and symmetry operations, the actual replacement of abstract tensors with explicit matrix representation represents a bottleneck for the computation. Indeed, complicated vertices can contain products of many gamma matrices to contract, which becomes quite expensive in practice. The current framework handles this problem by performing many simplifications on the level of the tensors and brute-forcing the replacement with matrices by dispatching the calculation on many cores simultaneously.

Before we describe the exact format of the ECRU module, one word on the convention of flavors and momenta in the description of vertex functions. For the Helac recursion the natural convention is to treat all flavors and momenta of input currents as incoming, while the flavor and momentum of the output current is outgoing. That ensures that the computed current can always be used as an incoming one in the next step of the recursion without any need to flip flavors or momentum directions. UFO models are structured to describe individual vertices with all momenta and flavors given as outgoing from the vertex. Thus, the interface must carefully perform the translation of conventions.

5.1.3 Technical description of the ECRU output files

ECRU stores the color-stripped currents alongside their color-flow configuration into an intermediate format. To illustrate the format we present the example of the SM $\bar{t}t \to g$ current in Figure 5.2. The current information is split into several blocks encapsulated by `<label>/</label>` markers. The current format includes four fields labelled `VertexData`, `ColorDecomposition`, `CurrentInternal` and `CurrentExternal`, whose layout we describe in the following:

- The `<VertexData>` block contains general vertex information. We will cover it in order, line by line. First, a unique label for this current is defined by `V_389_g_txt_1_2`. We derive it from the internal name inside the UFO model by extending it with the field content. The `1_2` identifies that this vertex contributes to two currents out of which this is the first one. The `2_2` current is generated by a SMEFT operator in this example. The label itself is arbitrary as long as each current has a unique one. The next line identifies the number of currents in the vertex, which is `3` in this example. The color-representation of the fields are given by `-3 3 8`, which encode the anti-fundamental, fundamental and adjoint representations respectively. Before specifying the flavors, we provide a mapping between the particle ordering in the UFO model and

the ordering within the HELAC current. By convention the first slot corresponds to the output current, while all others are input currents. In this case we specify `0 2 1 1 2 0`, which encodes pair-wise repositioning in the format $i_V \to i_J$, where i_V is the position in the UFO vertex and i_J the one the HELAC current. For this example, the particle 0 should be placed in slot 2, particle 1 in slot 1 and particle 2 in slot 0. We note that this mapping is not completely arbitrary. Interchange of input currents will finally correspond to interchanging the order of currents inside HELAC. This will result in the same expression for the current itself, but the interchange can generate a relative sign factor χ, which was introduced in Section 4.1.1. Squared amplitudes are blind to global sign factors, but whenever multiple vertices contribute to the same current one must make sure that both have correct relative sign. A straightforward way to achieve this, is to order the input currents according to the corresponding Lagrangian expression. This mapping should be applied to both the color representation and particle flavors, which are specified next. `-6 6 21` are the flavors of the involved currents using the original order inside the UFO model. We opt to use the standard PDG codes [5] at this stage, which differ from the internal ones inside HELAC. In line with the overall philosophy the translation is isolated to the HMODELGENERATOR program. In the present construction, ECRU will provide the flavors as outgoing from the vertex, while the momenta are already aligned with the expected convention in HELAC-NLO. Finally, the last two lines specify the perturbative coupling powers of the vertex. The first line specifies the powers and the second one the corresponding labels. For the EFT case we provide both a coupling power for the NP scale Λ as well as for each individual Wilson coefficient.

- The color-flow information is decoded in the `<ColorDecomposition>` block. Each line corresponds to a unique color-flow specified through Kronecker-deltas, which we denote by `cd(i{k},j{k})`. The indices correspond to the (anti-)color connection of particle `k` in the original UFO vertex ordering. Looking at the example file, we see that the first configuration corresponds to the regular gluon connection, while the second one is the self-connected $U(1)$ gluon. Note again that the HMODELGENERATOR program will convert the ordering of the labels `i{k}` and `j{k}` into the ordering of the actual currents using the mapping provided in the `VertexData` block.

- The `<CurrentsInternal>` block contains the actual four-component current expressions as a function of the input currents. Each line corresponds to one color-flow configuration. To increase readability we have inserted line continuations in the example, which are highlighted by `//`. The internal descriptor is used to declare that the current is contracted with the propagator of the output current flavor. We label momenta using `zp{k}`, fermions with `zf{k}`, anti-fermions with `za{k}` and finally scalars with `zs{k}`, where `k` labels the particle position in the UFO vertex. Unlike the mapping in `VertexData`, we are using 1-based indexing here instead of 0-based ones. Mapping the indices to the ones for the actual current is again relegated to the HMODELGENERATOR module. We note that the momenta are given in the HELAC convention,

[5] UFO models typically follow the PDG convention.

that is with input currents having incoming momenta and the output current having outgoing momentum, while the currents themselves are all given with outgoing flavors. The remaining translation of the flavor convention is performed in the HModelGenerator module.

- Finally, the `<CurrentsExternal>` block is identical to the previous one, but contains the expressions without any propagator contracted to the currents.

```
<VertexData>
V_389_g_txt_1_2
3
-3 3 8
0 2 1 1 2 0
-6 6 21
0 1 0 0 0 0 0 0
EW QCD NP ctp cpG ctG ctW ctZ
</VertexData>
<ColorDecomposition>
cd(i2, j3)*cd(j1, i3)
cd(i3, j3)*cd(j1, i2)
</ColorDecomposition>
<CurrentInternal>
-G*sqrt2*(za2_2*zf1_4 + za2_3*zf1_1) , //
-G*sqrt2*(za2_1*zf1_3 + za2_4*zf1_2) , //
G*sqrt2*(za2_1*zf1_4 - za2_3*zf1_2) ,  //
G*sqrt2*(za2_2*zf1_3 - za2_4*zf1_1)
G*sqrt2*(za2_2*zf1_4 + za2_3*zf1_1)/NCOL ,  //
G*sqrt2*(za2_1*zf1_3 + za2_4*zf1_2)/NCOL ,  //
G*sqrt2*(-za2_1*zf1_4 + za2_3*zf1_2)/NCOL , //
G*sqrt2*(-za2_2*zf1_3 + za2_4*zf1_1)/NCOL
</CurrentInternal>
<CurrentExternal>
-I*G*sqrt2*(za2_2*zf1_4 + za2_3*zf1_1) , //
-I*G*sqrt2*(za2_1*zf1_3 + za2_4*zf1_2) , //
I*G*sqrt2*(za2_1*zf1_4 - za2_3*zf1_2) ,  //
I*G*sqrt2*(za2_2*zf1_3 - za2_4*zf1_1)
I*G*sqrt2*(za2_2*zf1_4 + za2_3*zf1_1)/NCOL ,  //
I*G*sqrt2*(za2_1*zf1_3 + za2_4*zf1_2)/NCOL ,  //
I*G*sqrt2*(-za2_1*zf1_4 + za2_3*zf1_2)/NCOL , //
I*G*sqrt2*(-za2_2*zf1_3 + za2_4*zf1_1)/NCOL
</CurrentExternal>
```

Figure 5.2: *Example ECRU output file for the SM $\bar{t}t \to g$ current. To improve readability, line continuations* `//` *have been inserted. The actual format just contains each four-component current in a single line instead.*

The Helac program is organized into two phases. First, a skeleton generation phase is responsible for preparing the workflow of the recursion. The result is a series of instruction on how to combine the currents to obtain the final amplitude. Once this phase is completed, one can either immediately begin evaluating the matrix elements or store the skeleton into file for later use. The color-flow information together with the information on the included flavors and perturbative couplings is used to construct the vertex functions that drive the skeleton generation phase. Meanwhile, the currents can be directly translated to the numerical evaluation of

the currents, which is used during the actual Monte Carlo run. In other words, the `VertexData` and `ColorDecomposition` blocks contribute to the skeleton generation phase, while the `CurrentInternal` and `CurrentExternal` blocks govern the matrix element evaluation phase. The constants that enter inside the currents, such as `G` or `sqrt2`, are currently not automatically generated inside HELAC. Instead, we modify the code by hand to make these available. Lastly, we note that the particle content in SMEFT is identical to the SM by definition. As a result, we can translate all particle flavors naively using a simple map and do not need a more general translation layer at this stage.

5.1.4 Code generation with HModelGenerator

The second module of the interface is HMODELGENERATOR, which reads the provided ECRU output files and generates the required HELAC routines. The HELAC-TREE, HELAC-DIPOLES and HELAC-1LOOP require slightly different versions of the model data in practice, which can be generated using a unified run script. The internal workings of the module are straightforward. The current information is read and converted into HELAC routines using string operations, while collecting overall model information that is finally used to construct the embedding of these routines into the FORTRAN code. In this section we first present a short overview on the role that the generated files play.

- The current functions can be generated in two formats. The first one corresponds to the original HELAC structure. A file `pan2.f90` is generated that contains routines for all currents in the model. All translated conventions are already applied here. The file is embedded as an include file into the newly created `current` module. This approach works well at leading order, but comes with large compilation times at the 1-loop level due to the large number of generated routines and the potentially complicated formulas involved. To improve the compilation speed, we provide a second way of accessing these routines. To that end, we split the routines into multiple submodules that house only $\mathcal{O}(100)$ current functions each. These can then be compiled using multiple threads in parallel. The required `makefile` is also generated at this stage. We store the submodule files into a separate `currents` directory. The interface block required by FORTRAN is written to `intf_currents.f90`. Currently, only HELAC-1LOOP implements this approach, while the updated HELAC-TREE and HELAC-DIPOLES use the merged file instead.

- Vertex functions, which are used during the skeleton generation phase, are generated in a similar manner. First, the color-decomposition data is converted into array lookups on the internal representation of the color-connection. The flavor convention is adjusted so that the flavors of incoming currents is given as incoming rather than outgoing. Finally, the vertex function is generated, which includes the check on the perturbative powers of the input currents, selection of applicable color configuration, calculation of the number of connected internal $U(1)$ gluon lines, and insertion to the skeleton file when successful. Similar to the currents, we generate a `pan1.f90` file, which acts as an include file to a vertexfunction module. To reduce compilation time, we also provide a submodule structure, associated makefile and `intf_vertices.f90` file. Only

HELAC-1LOOP currently adopts this approach. The corresponding files are stored into a `vertices` subdirectory.

- The calls to the current routines are embedded into HELAC-NLO through switch statements, which are written to `currentcallsLO.f90` and `currentcallsUVandR2.f90` for LO and counterterms respectively.

- The vertex functions that govern the skeleton generation are embedded in two ways. The traditional way just includes a blind test of all vertex functions applicable for a combination of `n` currents. The calls to these are written to four files `include_{type}_{n}particles.f`, where `type` is given by `tree`, `r2`, `loop_epsm0`, `loop_epsm1` and `n` corresponds to the number of currents. In practice, this implementation is only used for debugging purposes. We have implemented a pre-selection algorithm[6] that obtains a list of potentially applicable vertex functions based on the previous history of the recursion. The actual selection then happens through a switch statement that is written to `vertexselection_{type}.f90`, where `type` can be `LO`, `UV` or `R2`. The separation of the finite and pole parts of the UV counterterms is handled by the pre-selection algorithm.

- The perturbative powers associated with each vertex are integrated via the generated `include_commons.h` include file. In addition, we include an on/off switch for all vertices, which can be used to disable the Higgs contributions according to the selection in the HELAC-NLO input files. A second file, `include_physics.f90` contains the calls to modify the latter at runtime.

- For SMEFT models we automatically provide a `Wilsonnames.h` include file, which contains a string label for each Wilson coefficient. In practice, it is only used to improve the optics of the output files in HELAC-NLO. However, it can also be used to quickly confirm the identity of the generated model.

- Several files containing global information on the model are needed to support the previously mentioned filtering algorithm during the skeleton generation phase. For each type of vertex, `LO`, `R2` and pole/finite pieces of `UV` counterterms we write five files. The algorithm is explained in detail in Section 5.3, so we just briefly sketch the role of each file. One include file to define array dimensions `arraysizes.h`, which correlates with the number of vertices in the model. `searchtable.h` includes the definition of the array used in the search algorithm. The `ptrarray.h` include file provides the map to potentially multiple vertices that can match a given set of input flavors and `vertexids.h` provides the actual lookup table for the vertex ID used to call the corresponding routine. Lastly, information on the output flavor of a given vertex function is made accessible through `outflavors.h`.

The full set of generated files is stored in a `model` directory, which can be placed inside the HELAC program codes and compiled using the included `makefile`. One additional file inside the HELAC program requires modification to complete the interface. Inside the `config` subdirectory, one configures the `SMEFT_configs.h` file,

[6] Discussed separately in Section 5.3.

which is currently not automatically provided together with the model routines. The parameters of the file are as follows:

- `N_operators` is the number of operators in the model.
- `NumberOperatorsInME` is the maximal number of operators that are to be included on the level of the amplitude.
- `NumberOperatorsInXS` is the maximal number of operators that can enter the squared matrix element.
- `LambdaScale` is the NP physics scale Λ that is included in the current definition of the model files.
- `WilsonLambdaScale` is an array of size `N_operators` that contains the Wilson coefficients at the scale `LambdaScale`.
- `NPpowers` is an optional array of length `N_operators` that specifies the NP powers of each operator. For SMEFT at dimension-6 all values are set to 2, which corresponds to Λ^{-2}.
- `checkNPpowers` is a flag that performs an optional check on the NP power of the amplitude during the recursion. When all operators share the same NP powers, one can safely disable this feature. In that case `NumberOperatorsInME` is sufficient to fully specify the included terms. When enabled, the information from `NPpowers` is used to perform this check.
- `ipowNPmaxinput` is the input maximum NP power to be used, when activating `checkNPpowers`. For example, when working up to Λ^{-4} this check would allow a contribution of a dimension-8 operator interfering with the SM amplitude and a dimension-6 squared contribution, but a dimension-8 squared contribution would be removed.
- Finally, `MaxPartInVertex` defines the maximum number of particles that can be present in a given vertex for the model under consideration. Note that this can be less than the operator mass dimension of depending their field content. For example, four-fermion operators enter at dimension-6, but only contain four fields. Lowering this value to the minimum is recommended, since several code paths can be terminated early based on this information.

An example `SMEFT_config.h` input file is shown in Figure 5.3 for the case of five dimension-6 operators. The input file itself is written in generic Fortran. The Helac-Smeft framework defines its state at compile time based on the information given in the `SMEFT_config.h` input file. Setting either `NumberOperatorsInME` or `NumberOperatorsInXS` to 0 forbids any SMEFT contribution. The code can detect this during compilation and adjust array sizes as well as omit code branches that only contribute to the SMEFT case. It is convenient to set `N_operators = 1`, when using the code in SM mode. The reason is that Fortran does not allow static arrays of length 0. To avoid this limitation, one would have to remove the declarations of arrays of size `N_operators` when compiling for the SM case. In practice, most arrays are defined from derived variables to avoid this problem, but the expressions in the input file are more transparent to the user when defined in terms of `N_operators`.

While the SM prediction can be generated this way even when using a SMEFT model, one should still employ an actual SM model in practical calculations. The actual SM generates significantly less vertex and current functions, which improves the compile time and eliminates unnecessary checks on the excluded SMEFT vertices during the skeleton generation.

```
!*****************************************************************************
!...
!... USER INPUT for SMEFThandler
!... Please make sure that this code contains valid Fortran 90 syntax.
!... These variables are included in the SMEFThandler module that can
!... be found in src/smefthandler.f90
!...
!*****************************************************************************

!... Number of SMEFT operators
!... [ Don't set this to 0 as some arrays use it to set their size. Use 1 instead. ]
integer, public,  parameter        :: N_operators           = 5

!... Number of SMEFT operators inserted into the Amplitude
!... [Setting to zero will only compute SM]
integer, public,  parameter        :: NumberOperatorsInME   = 1

!... Number of SMEFT operators inserted into the cross-section
!... [Setting to zero will only compute SM]
integer, public,  parameter        :: NumberOperatorsInXS   = 2

!... Scale of the UV theory in GeV
real(kind(1d0)), public, parameter :: LambdaScale           = 1000

!... Size of Wilson coefficients at scale LambdaScale
real(kind(1d0)), public, parameter :: WilsonLambdaScale(N_operators) = (/1,1,1,1,1/)

!... Inverse powers of Lambda for each operator
integer, private, parameter        :: NPpowers(N_operators)          = (/2,2,2,2,2/)

!... Explicit check on NP powers is only required, when
!... the result does not allow all insertions defined
!... above due to constrains on the NP power
logical, public,  parameter        :: checkNPpowers         = .false.
integer, public,  parameter        :: ipowNPmaxinput        = -1

!... Maximum number of particles in the vertexfunctions
integer, private, parameter        :: MaxPartInVertex       = 6
```

Figure 5.3: *Example* `SMEFT_config.h` *input file for a SMEFT setup with 5 operators and only linear contributions at the amplitude level. The NP scale* Λ *is set to* 1 TeV*.*

5.2 Implementing SMEFT in Helac-1Loop

A general overhaul of all three codes of the Helac-Nlo framework is needed to facilitate the computation of SMEFT predictions. While the code bases of Helac-Tree, Helac-1Loop and Helac-Dipoles contain many overlapping modules, they have largely not been kept synchronized, and thus require dedicated modifications for all three. For Helac-Smeft we initially built a prototype in Helac-Tree and then put most attention towards the most complicated case, which is Helac-1Loop. The modules from Helac-1Loop were then simplified and finally embedded in Helac-Dipoles. This increases the interoperability, although the inherent restrictions of the code base cannot fully be avoided. In this section, we

describe the general updates needed and their respective implementations into the HELAC-1LOOP code. For those familiar with the HELAC program code, we will make explicit reference to the files that were modified, removed or newly created for this project.

In accordance with the goals for the project, we want to book-keep several SMEFT amplitudes simultaneously. In practice, this means that simply keeping track of powers in the NP scale Λ is not sufficient. Instead, we want to introduce a new integer label that corresponds to the contained Wilson coefficients of a current. The number of required labels depends on the number of included SMEFT operators. When no operator is included the code should seamlessly revert to the original logic. We first focus on the case of only one operator insertion in each amplitude and squared contributions to cross sections. The extension to the fully general case is straightforward. In this simplest case, we require

$$\texttt{N_IDs} = (\texttt{N_operators} + 1)$$

labels for linear contributions [7]. By convention 0 corresponds to the SM term. There are

$$\texttt{N_IDs_XS = (N_operators + 1)(N_operators + 2)/2}$$

combinations for squared contributions, where we organize the bookkeeping purely by powers of Wilson coefficients, identifying $C_i \sim C_i^*$ and combining conjugate products $C_i C_j^* + C_i^* C_j$. The implementation is provided by the newly written `smefthandler.f90` module, which defines the number of labels at compile time and provides the routines to map the combination of labels through a `calculateID` subroutine[8]. The variables such as the NP scale Λ and the numerical values of the Wilson coefficients are also contained within. In addition, a `getOperatorName(s)` allows for convenient access to a string representation of a given label. By including the module into other parts of the program, we can compile the size of arrays based on the number of labels required. This allows the code to revert to the SM implementation at compile time, which increases the maintainability of the program.

The inclusion of vertices with more than four legs requires the extension of several arrays inside the HELAC code base. Each step of the Dyson-Schwinger recursion is stored in an array. The SM algorithm uses 18 elements for that purpose, which includes both the current information and additional information on the vertex itself. As previously mentioned, we have replaced the way colors are stored. The original code provided just one integer per step, while the new one requires two. Fortunately, we can repurpose the previously used helicity index, which is not strictly needed. The two additional currents then require an additional 6 slots. The array tracking the perturbative powers is also extended at compile time to house one additional integer label and the additional currents, which appear in the EFT calculation. We use the state of `smefthandler.f90` to define the array dimensions at compile time. The `pan1.f` file that contains the logic for skeletons and their evaluation was refactored to simplify the development process. All vertex functions are moved into a

[7] In practice, it is more convenient to build a more complicated logic that takes the `NumberOperatorsInME/NumberOperatorsInXS` variables into account.

[8] The current logic to organize cross-sections is straightforward, because we compute all contributions at $\mathcal{O}(\alpha^n \alpha_s^m)$. For the extension to the EW sector one would have to introduce separate book-keeping functions for the cross section. Such an extension can be easily achieved and is planned in the near future.

separate `vertexfunctions.f90` module. The `pan2.f` file that previously contained the current function definitions is replaced by `current.f90` module instead. The remaining logic for evaluating the matrix elements based on the skeletons is now contained in a renamed `pan.F90` module file. Extensive adjustments have been made for that purpose. First, the evaluation functions for the LO and 1-loop numerators are updated to support the additional currents. Next, the 1-loop amplitude reduction is automatically repeated for each operator amplitude separately. To improve efficiency, we do not recompute the numerator function for each reduction. Instead, we evaluate the full recursion with all operators during the reduction of the first loop amplitude, which yields an array of numerators corresponding to each Wilson coefficient. These can be used directly for the reduction of the remaining 1-loop amplitudes without reevaluating the numerators separately. This approach effectively enables the complete reuse of all SM sub-currents in the recursion of the SMEFT numerators, while also scaling gracefully to the Born-like computations at LO and in the real-emission processes. This approach necessitates that all amplitudes are computed with the same cut-off value for the integral rank in CutTools, because the probed momenta in the OPP method can be misaligned otherwise. However, we currently don't implement any automatic detection of the integral rank for the general EFT, so we must always use the maximum possible integral rank throughout all reductions anyways. An additional check is implemented to only repeat the reduction for those amplitudes, whose corresponding numerators are genuinely non-zero.

To improve the runtime performance under the additional complexity of the SMEFT calculation, we implement two additional optimizations for the evaluation of 1-loop numerators. When constructing the numerator function of a specific loop topology one can categorize the sub-currents into two groups. First, currents that contribute to the loop structure itself and thus depend on the explicit loop-momentum q, which is requested by CutTools. Second, currents that are outside the loop structure and are unaffected by the loop momentum. The distinction can be trivially detected in Helac-1Loop through the binary momentum index of the current. All currents with $m \leq (2^n - 1)$ are LO-like, while the ones with $m \geq 2^n$ explicitly depend on the loop momentum. It would be most efficient to pre-compute all the loop momentum independent currents for all numerators in the beginning, and simply compute the additional q dependence during the reduction itself. However, the current structure of the code makes this approach difficult to implement. Instead, we restrict ourselves to individual numerators, where this recycling structure is more straightforward to implement. Practically, we fill the LO-like currents to a separate buffer, which we use to reload them at each evaluation of the numerator function. Note that this does not introduce an additional cost over the regular approach, because the array storing the recursion needs to be initialized in either case. This less ambitious approach still yields a significant improvement in the runtime of the code without requiring extensive restructuring of the 1-loop skeleton. Secondly, the usage of modern numerical reduction techniques, such as the OPP method, requires careful handling of numerical instabilities. In Helac-1Loop we implement two approaches:

1. If the partonic process contains a gluon then we test the Ward identity by replacing the gluon polarisation vector with its respective momentum [9]. Ana-

[9] In practice, its optimal to rescale the momenta by the energy scale of the process. This avoids

lytically one then expects each color-stripped amplitude to vanish separately. To test this expectation we sum up the absolute values of the leading-order, finite 1-loop and singular contributions at the amplitude level and compare against an predefined numerical threshold [10]. If the threshold is exceeded, we store away the phase-space point separately and recompute it with higher floating point precision.

2. If no gluon is available inside the partonic process, then we perform a rescaling test as outlined in Ref. [228]. Here one uses information on the mass-dimension of the Lagrangian and amplitude to relate the amplitudes, when all massive quantities are rescaled by a common factor. One can then compare the expectation with the numerical result to obtain a stability indicator.

Generally, processes with gluons are more expensive to evaluate. Therefore, we focus on improvements to the former case. First of all, we perform the reduction of the 1-loop amplitude for the gauge invariance check on the sum of the SM and SMEFT contributions to avoid the additional cost of individual reductions for each amplitude. The cost of the test can be reduced further by only recomputing sub-currents that contain the selected gluon. This can be accomplished by performing a bitwise `AND` operation on the binary momentum integer of the gluon and the sub-current under consideration. If the result is 0 then the sub-current does not need to be recomputed. For a gluon at the endpoint of the recursion, we can skip straight to the end and don't recompute the recursion. Only the final contraction as well as the reduction have to be repeated in this case. Although this does not accelerate the performance of the code in a typical one-point benchmark, we find significant gains in real-world scenarios. The implementation requires adjustments to the matrix-element evaluation in `pan.F90` of course and changes to the `stabilitycheck.f90` file, which implements all stability checks. The previous implementation of the gluon Ward identity check instead fully recomputed the amplitude which effectively doubles the cost for each phase-space point. Another improvement could be achieved by dynamically adjusting the endpoint of the recursion to the first gluon inside the partonic process. We currently don't implement such an optimization due to constraints of the original HELAC implementation.

The role of the `physics.f90` file is redefined in the new framework. Previously, this file would implement the calculation of coupling strengths for each vertex and update couplings whenever the strong coupling changed due to the running of α_s. In the new framework, we directly multiply all constants inside the current functions, which eliminates the need to update coupling strengths. Instead, the currents now draw all constants from a common buffer that allows direct modification as needed. The remaining functionality of the file now contains the evaluation vertex switches that can remove specific contributions of the model, including a general removal of Higgs contributions. Furthermore, it acts as a controller of several other actions, including the computation of complex masses, running Wilson coefficents if enabled and adjusting massive quantities for the scale-test.

different behaviour based on the momentum fractions of the supplied phase-space point.

[10] Reasonable thresholds depend on the complexity of the integrand. In practice, values between 10^{-5} and 10^{-10} work well for double-precision calculations.

The LO recursion is implemented inside `tree.F`, which was extended to perform combinations of four and five input currents. The implementation inside HELAC always selects the momentum of the output current first and then constructs all combinations of input momenta that can produce it. The underlying routines that split the binary momentum have been reimplemented inside the `intpar.f` file. The tree-recursion was further modified to support the flavor selection algorithm as described in Section 5.3. The implementation of the NWA has to be generalized in several ways for SMEFT calculations. We will outline changes alongside other adjustments, when they come up. The first is that the compatibility of a combination of currents in the LO recursion with the decay chain in the NWA has to be tested now also for five- and six-particle vertices. To that end, the `helac_decay_match` routine inside `decay.f90` was generalized for an arbitrary number of input currents. The last adjustment to the LO recursion relates to the computation of symmetry factors of the given process. The original implementation was tailored to the calculation of off-shell amplitudes, which necessitated by hand adjustments for processes in the NWA. A general implementation has been included in `intpar.f`, which now correctly computes the symmetry factor for each factorized decay separately and returns the correct factor for the total process in the end.

The situation on the level of recursion for the 1-loop amplitude is more complicated. The technical implementation is provided in `loop.F`. The additional LO-like routines that handle the construction of counterterms and R2 terms have to be adjusted in the same way as the LO recursion. We take this opportunity and update the underlying renormalization logic. The by-hand construction in the original version of HELAC split the renormalization into several pieces. Wave-function and α_s renormalization were implemented using the computed LO amplitude and only then combined with the 1-loop amplitude. Meanwhile, the calculation of the mass renormalization had to be included inside a dedicated LO-like recursion, just like the R2 terms. UFO models are constructed to include the coupling and wave function renormalization as dedicated counterterms, which prompts us to remove the old separate treatment and instead include all renormalization counterterms in the same vein as the R2 terms. This removes the need for the `renormalization.f` file, which housed the previous logic and unifies everything into the same framework as implemented in `loop.F`. Several revisions are needed in the construction of the 1-loop topologies. While the first adjustment is needed only due to performance consideration, we also have to make structural adjustments to properly include all SMEFT contributions. Starting with the performance optimizations, we focus on the flavor dressing of the generated loop topologies. Once a loop topology is generated, we must dress the loop propagators with flavors and colors to ensure consistent propagator masses as required by the OPP method during the reduction. Once a configuration is selected, one can in principle simply start generating the recursion relation using the regular HELAC algorithm. In practice, naively looping over flavors and colors quickly becomes very expensive. It is much more efficient to first reduce the number of topologies using the constraints of the model. In the original HELAC-NLO a dedicated test of the quantum numbers of the LO currents that enter the loop was performed, which could then restrict the possible flavors of the loop-propagators. This approach doesn't generalize to arbitrary models and was removed for that reason. In its place, we perform two new tests instead.

1. Since all the currents entering the loop are LO-like, we can build a set of allowed currents at LO and then directly check the compatibility with the loop topologies. To that end, we have to perform the LO recursion also for the currents containing the endpoint momentum. The computational cost of doing so is negligible.

2. The possible flavors of LO currents entering the loops can be directly used to restrict the possible vertices at each vertex inside the loop. Two neighbouring propagators must always connect to one or more LO-like currents [11]. We can use the same LO filtering algorithm described in Section 5.3 to search for valid contributions.

Existing filters that enforce the color-conservation on the level of the loop topology can be reused without modification, because the SMEFT by definition only contains SM field content. However, a generalization would be much easier to implement, because knowledge of the color-representation of each field is already fully sufficient to find invalid color connection configurations.

Typical calculations inside HELAC-NLO require the renormalization of massive particles. The choice of the on-shell scheme is very attractive for this purpose as it avoids the computation of on-shell 2-point functions[12] on the external legs or alternatively the inclusion of Z-factors from the LSZ formula [102]. When treating top-quarks (or other colored massive partons) in the NWA one must similarly make sure to not calculate the self-energy contributions on the internal top lines. The formulation of the loop topology generation in HELAC-NLO is not well suited to this. To understand this point recall that a given loop topology is initially characterized by the momenta that are inserted into the loop through LO-like currents. Note that this formulation inherently cannot distinguish the number of external legs that connect to each vertex inside the loop. A schematic example is provided in Figure 5.4, where the momentum i_1 can be connected to a three- or four-particle vertex. We will see that this property of the algorithm becomes quite problematic when applying the on-shell scheme to the calculation. Consider the two loop contributions to the external legs in Figure 5.5. The left diagram is part of the self-energy correction to the propagator and should not be computed in the on-shell scheme. However, the right diagram can contribute for specific SMEFT operators and is not part of the self-energy. Note that the momentum representation used to construct the loop topologies is identical for these two diagrams. Consequently, disentangling these contributions is not possible during the construction of loop topologies and instead must be moved to the Dyson-Schwinger recursion. The original implementation of HELAC-NLO does not implement this at the level of the recursion however. Instead, all bubble contributions that connect to an external leg are simply dropped. This leads to an omission of a diagram similar to the right one, where the external top is replaced with a gluon. Conveniently, such a diagram is scaleless for massless gluons, which leads to a vanishing contribution in dimensional regularisation. In SMEFT such contributions can appear with massive external legs, which lead to non-vanishing

[11] Multiple LO currents can be connected due to the existence of 4-, 5-, and 6-particle vertices.

[12] Note that on-shell self-energy contributions for massless particles vanish in the $\overline{\text{MS}}$ scheme as well, which unifies the treatment of both.

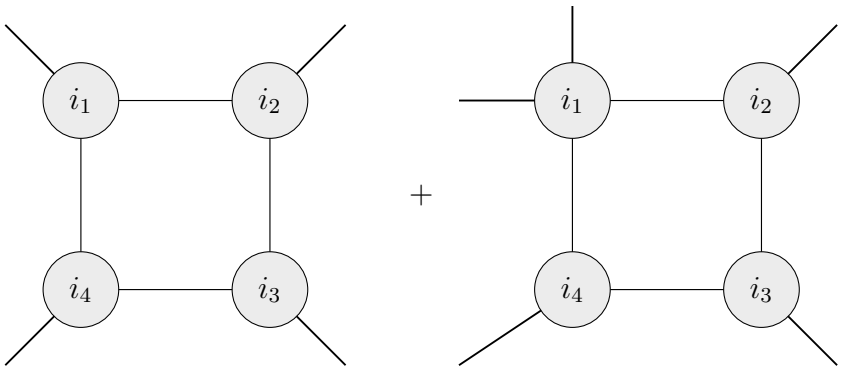


Figure 5.4: *Loop-topologies with identical momentum insertions contain contributions with different number of external legs. The provided example shows a box-topology with either 3 or 4 particle vertex connected to the momentum insertion i_1.*

terms [13]. Generally, the number of propagators in a loop does not correspond to the number of legs of the loop function. This discrepancy is the underlying issue in the original implementation. The diagram showcased in Figure 5.5 can still be handled

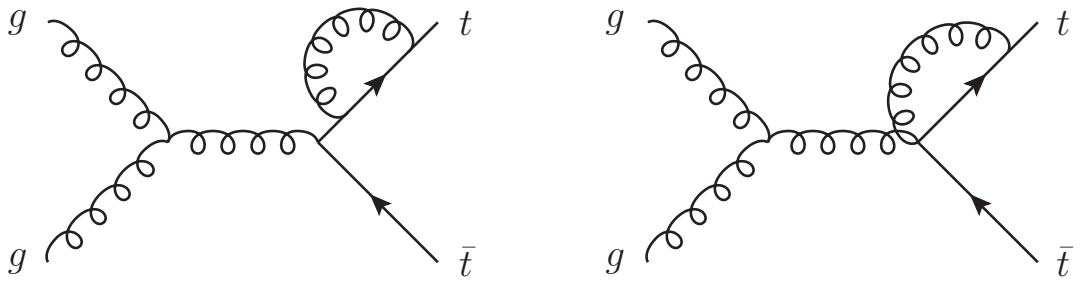


Figure 5.5: *Shown are two bubble loop contributions to the external legs in $gg \to t\bar{t}$. The left diagrams contributes to the self-energy and is exactly cancelled by the renormalization conditions in the on-shell scheme. The right contribution can be induced by SMEFT operators, but does not represent a self-energy correction.*

on the level of the recursion. We adjust the loop topology generation so that neither diagram is pre-emptively removed. To subsequently identify such diagrams during the recursion we make the following observations:

1. All self-energy contributions are generated from two three-particle vertices.
2. The $(n+2)$-particle recursion, used to construct the loop-numerator, always combines currents from the external leg inwards. This can be understood by noticing the correspondence of the level of the current and the direction of the recursion in the diagram.
3. The vertex that closes the loop is given in the recursion by the vertex, which takes both extra-particles as a part of the input currents. The level of such a current is always exactly three, accounting for one external leg as well.

It is easy to detect such a vertex in the recursion. One first checks the level of generated current and for the presence of both extra-particles. If the level of the current

[13] Such topologies would also contribute to a NLO EW calculation due to the massive W and Z bosons.

is 3, then we are always looking at a self-energy contribution and should reject such a contribution. Otherwise, we are looking at the second kind of diagram and should include it. This prescription works for all legs except at the endpoint of the recursion, where one must slightly adjust the logic. We instead apply a restriction on the second to last vertex, which can again be detected using the level of the generated current. By confirming that only one of the extra-momenta is contained within the generated current, we automatically know that the last vertex can only connect to three legs and contains the second extra-particle. Consequently, the current (i.e. second to last vertex) shall not be generated by a 3-particle vertex as well. For external legs this algorithm is fully sufficient and has been implemented inside the HELAC-SMEFT framework. The treatment of colored particles in the NWA adds additional complications. As with the external legs, one must remove the self-energy loop diagram from the calculation. One could be tempted to apply the same algorithm as outlined above. However, one must include two configurations in this case as is shown in Figure 5.6. It turns out that the algorithm outlined above can-

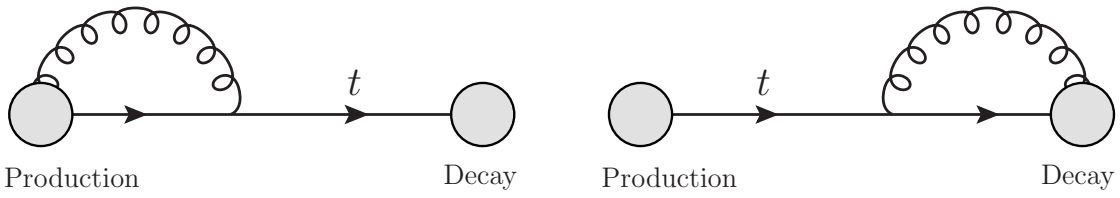


Figure 5.6: *Non-vanishing 2-point loop contributions can connect to on-shell top-quark lines in two configurations. The gray blobs represent the production and decay processes in the NWA.*

not resolve both contributions correctly, while also removing the genuine self-energy correction to the on-shell top-quark. Before we have used the fact that all 2-point loop integrals that should be included include at least one vertex with more than 3 particles and that we can predict the order in which they appear in the recursion. It should be clear that the position of the 3-particle vertex is exactly swapped in the two configurations in Figure 5.6. The solution to this problem is to decide a priori, which contribution one wants to generate and then apply the corresponding condition separately. We currently do not solve this issue fully. Instead, we handle the direction with the gluon loop connecting to the production correctly and defer the full solution to future work. These corrections are implemented in `loop.F` and are complete for all off-shell processes and for the application with top-quarks in the NWA as long as the decay does not include a gluon at LO. We further note that the SM case at NLO QCD is actually covered by the existing implementation, because the only four-particle vertex that enters loops is the four-gluon vertex. However, EW loop corrections also require a generalization in the same scenario.

The construction of fully reweightable LHE files in SMEFT is done in such a way that they natively collapse back into the previously implemented SM case when no SMEFT operators are present. However, the actual implementation was fully replaced for this project to better deal with the complexity of the underlying code. The extraction of the logarithmic dependence in the 1-loop calculation is implemented as described in Section 4.4. The counterterms in SMEFT can have an explicit dependence on an additional scale μ_{EFT}. Since the CPU time required

to compute CTs is negligible compared to the unrenormalized 1-loop amplitude, we simply perform the calculation at multiple scales and solve the system of equations to obtain a decomposition as

$$w_{\rm CT} = c_{\rm L00} + c_{\rm L10}\log(\mu_R^2) + c_{\rm L01}\log(\mu_{EFT}^2)\,. \tag{5.8}$$

Together with the unrenormalized 1-loop amplitude we then arrive at a decomposition into four coefficients $c^{(i)}_{\rm L00}$, $c^{(i)}_{\rm L10}$, $c^{(i)}_{\rm L20}$ and $c^{(i)}_{\rm L01}$, where the two indices represent the power of the logarithms in μ_R^2 and μ_{EFT}^2 respectively. The index i corresponds to each of the partial contributions. The coefficients have to be extracted for each partial SMEFT weight separately to conform with the overall goal of producing independent SMEFT cross sections for all operators at the same time. For completeness, we report the additional information stored in the LHEF. The `init` block is augmented by additional information on the included operators, physical parameters entering their running matrix [14], the Wilson coefficients at the NP scale Λ and the mapping between the included coefficients and the ordering of their respective weights. Each `event` block now stores the coefficients $c^{(i)}$ for each SMEFT contribution, the Wilson coefficients in this specific event as well as the scale μ_{EFT}. This information is sufficient to fully reconstruct the dependence on Λ, μ_{EFT} and all Wilson coefficients. The implementation is included in `LHEFIOhandler.f90` and is called by a newly written `LHEFreweight.f90` module. HELAC-1LOOP always performs the integration using the previously described reweighting technique. As a result, it is the only program among the HELAC codes that does not produce the full LHEF on its own and instead can copy several blocks directly from the input LHEF.

The HELAC-1LOOP program is designed to reweight leading order LHEF generated by HELAC-TREE, which has also been modified using many of the same principles presented here. The only part that exceeds the capabilities of HELAC-1LOOP is the explicit interface to LHAPDF, the interface to Monte Carlo and phase-space optimization modules and the routines to generate LHEF, which have been minimally expanded for this purpose. The organisation of the recursion and matrix-element evaluation is otherwise the same as HELAC-1LOOP. However, many improvements to the code quality have not been introduced to the HELAC-TREE code, because the much smaller project size did not induce a large pressure to refactor the program. A long-term perspective would be to integrate the missing functionality into HELAC-1LOOP and reduce the size of the code base to only two codes instead.

The calculation of the spurious terms and R1 contribution in the OPP method requires knowledge of the functional form of their loop momentum dependence. For renormalizable theories this can be constructed once and for all by exploiting that any loop integral can be written in a renormalizable gauge, which bounds the number of loop momenta in the numerator by the number of propagators [201]. This assumption is no longer fulfilled for effective theories, where the rank of the integrand can exceed the number of propagators. Updated versions of the CUTTOOLS program do support calculations where the rank exceeds the number of propagators by at most one. In the context of the $pp \to t\bar{t}H$ process the $\tilde{\mathcal{O}}_{\phi G} = \left|\phi^\dagger\phi\right| G^a_{\mu\nu}G^{a\mu\nu}$, where $G^a_{\mu\nu}$ is the gluonic field-strength tensor and ϕ the Higgs doublet, is one such operator

[14] These parameters are needed to apply the running of the Wilson coefficients retroactively in an automated way.

that can generate high integral ranks. Figure 5.7 shows two such contributions. Consequently, we replace our previously interfaced version `1.2` of CutTools with

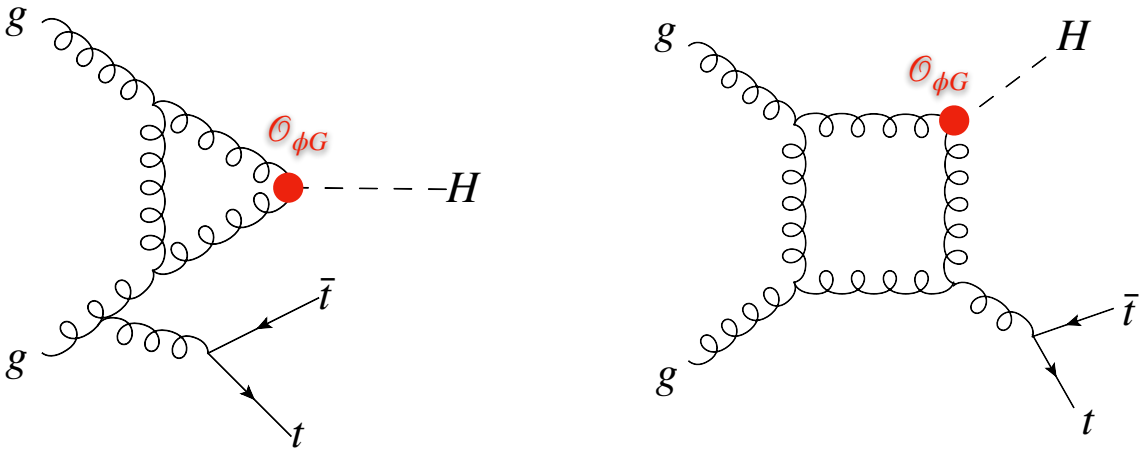


Figure 5.7: *Examples of loop integrals with rank exceeding the number of external legs. The SMEFT operator $\mathcal{O}_{\phi G}$ induces gluon Higgs couplings that are important to study in the $pp \to t\bar{t}H$ process. Taken from Ref. [3].*

a newer version `1.9.3`. One unfortunate side-effect is that the newer versions of CutTools are distributed with the very outdated version `3.4` of the scalar loop library OneLoop, which is indeed older than the previously interfaced version `3.6`. Important for the project in this thesis is the observation that the computation of $pp \to t\bar{t}H$ in the NWA in SMEFT indeed runs into bugs within the scalar loop library when using this setup. Our observed problems are indeed resolved after interfacing the newest version of OneLoop (`v3.7.2` at the time of writing). As a bonus, additional numerical stability improvements are also included in the updated code.

5.3 Optimization of the LO skeleton generation

The Dyson-Schwinger recursive algorithm combines currents based on their contained momenta, indicated by the binary indices i_j. However, in models with more than one field one generally has to test the currents for their respective flavors, colors, and contained perturbative coupling powers to determine if a combination should be included in the recursion or not. In models with many vertices this quickly becomes very expensive as many vertex functions have to be executed. The original implementation of Helac-Nlo optimizes this process by computing select quantum numbers for each current of momentum i, where we use the binary representation native to the Helac program. To that end, one can simply combine the quantum numbers of the contained external currents. For example, the SM does not contain a vertex that combines a quark and lepton. Similarly, the loop topologies are pre-filtered using a similar logic. In this case, the code is even more specific, often making use of restrictions that only appear at NLO QCD in the SM. To generalize the code to potentially arbitrary EFTs this logic had to be removed, which massively slows down the underlying code. This necessitates the introduction of a replacement algorithm. To make the inclusion seamless several key properties should be respected:

- The algorithm should select a subset of vertex functions based on the momentum of the input currents under consideration.
- A straightforward extension to the filtering of loop topologies should be possible.
- The strategy must be applicable to any model and thus not use explicit physics knowledge.

In order to accelerate the recursion, we make use of the global model information. Given any set of input flavors, we already know all vertices that can be applied and simply directly sample these vertices. To that end, we store the flavors of generated currents of a given momentum in an separate array. When the algorithm tries to combine currents with the set of binary momentum indices $\{i_k\}$, we can directly query all the flavors $m_k^{(j)}$, where j labels the different flavors available for the momentum i_k. Next, we test all sets of flavors for the input currents $\{m_k^{(j_\sigma)}\}$, where σ is used to label the possible combinations. Any of these sets of flavors can be mapped onto a unique integer using

$$I(m_1, ..., m_n) = \sum_{k=1}^{n} 64^{k-1} \cdot (m_k + m_0) \,, \tag{5.9}$$

where $m_0 = 14$ is an offset used to make each contribution positive [15]. For vertices with at most 5 input currents one can fit these integers I into a 4-byte integer type. Note that this unique integer can be pre-computed for each vertex in the model in the same way as we compute it for our candidate current candidates combinations. The task thus reduces to efficiently checking the integer I_C of the current against all I_V in the model. While generating the model data, we fill all unique vertex IDs into a list. Of course, more than one vertex can share the same ID. For example, $\bar{q}q \to \gamma$ and $\bar{q}q \to Z$ both share the same input currents and thus correspond to the same integer I_V. We keep a second list that stores the actual vertex IDs that correspond to a given I_V. Once the model is fully parsed, we then sort the first list. To map the indices of the first list to vertex IDs we make use of an array of pointers. The vertex IDs are written into an array, while ensuring that vertices with identical I_V are placed sequentially. Each element of the array of pointers then points towards a subset of vertex IDs corresponding to its index. Similarly, we prepare a second array of output flavors that plays the role of the array of vertex IDs. To summarize we prepare four arrays

$$\begin{aligned} \text{searchtable} &= (I_1, ..., I_n) \,, \\ \text{ptrarray} &= (p_1, ..., p_n) \,, \\ \text{vertexIDs} &= (v_1, ..., v_m) \,, \\ \text{outflavors} &= (f_1, ..., f_m) \,, \end{aligned} \tag{5.10}$$

where m is the number of vertex functions in the model and n is the number of vertex functions with unique input flavors, i.e. $m \geq n$. We then search it with the following three steps:

[15] In HELAC convention flavors are labelled from -13 to 41. If the flavor convention exceeded 64 used slots then we would adjust the base-64 to a higher number and pay the cost of using a larger data type.

1. Search the sorted `searchtable` using standard binary-search algorithm [16]
2. If the first step is successful, we select the pointer located at `ptrarray(ind)`, where `ind` is the index at which we found the match in step 1.
3. Assign a public pointer to the pointer from step 2, which acts as the point of retrieval for the user.

The logic can be applied to the flavors in the exact same way. Only the corresponding `ptrarray` is provided separately and now points to the `outflavors` array. The Dyson-Schwinger recursion can now simply loop through the slice that the public pointer is assigned to, to retrieve all vertex IDs of the vertex functions that should be called. Note that some or all of the vertex functions might still return without adding any entry to the skeleton, because we have not checked any color or coupling power constraints at this point. We compare the performance during the skeleton generation phase against the original implementation using quantum numbers for several SM processes. The results were obtained on a machine with a Intel(R) Core(TM) i5-14600 processor and the `gfortran` compiler included in version 7.5 of the `gcc`. Both codes were compiled with the `-O3` optimization option. Table 5.2 shows the comparison for several off-shell matrix elements. Many processes are accelerated by a factor 10 or more. This highlights the drastic advantage of this strategy for the leading-order recursion. This approach is used for the recursion algorithm in all three codes, Helac-Tree, Helac-1Loop and Helac-Dipoles. However, the Helac-1Loop code employs the same routines to filter loop topologies. The details are discussed in Chapter 5. The technical implementation is provided in the newly

Tree-Level Skeleton - CPU Time (s)			
Process (SM)	Helac-Smeft	Helac-Nlo	Ratio
$gg \to ggg$	0.011	0.173	0.06
$gg \to gggg$	0.070	5.760	0.01
$gg \to ggggg$	2.574	222.410	0.01
$gg \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}$	0.016	0.069	0.23
$gg \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H$	0.030	0.286	0.10
$gg \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,g$	0.064	0.786	0.08
$gg \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,gg$	1.016	24.094	0.04

Table 5.2: *Performance benchmark of the Born-level skeleton generation for various SM processes. Compared are the updated Helac-Smeft framework and the Helac-Nlo program. Off-shell matrix elements are considered in all cases and the programs have been compiled using the `gcc-7.5` compiler with the `-O3` optimization flag enabled. The computer is equipped with an Intel(R) Core(TM) i5-14600 processor. Taken from Ref. [3].*

written `skeleton_helper.f90` module for all three codes. After an initialization call to `init_skeleton_helper` and `init_generatedFlavors` one can automatically store all generated flavors during the recursion using the `update_generatedFlavors` routine. The unique integer I can be retrieved using `setInputID` and finally one

[16] The binary search algorithm scales as $\mathcal{O}(\log n)$, which ensures good scaling behaviour even for very large number of functions.

can set the public pointer to the slice of vertex IDs using the `selectVertices` subroutine. Functionally, the HELAC-DIPOLES is identical to HELAC-TREE for this purpose. The adjustments of the module in HELAC-1LOOP are minimal. We simply introduce separate routines to select vertices that are part of the tree-level Lagrangian, UV counterterms or R2 terms using the identical logic of the LO application. Additionally, we provide a routine to retrieve the output flavors of the vertices through a `selectFlavors` subroutine.

5.4 Implementing SMEFT in Helac-Dipoles

A previous implementation of SMEFT in HELAC-DIPOLES was developed for a selected set of operators in Ref. [229]. However, the implementation was done by hand and using an earlier version of the HELAC-DIPOLES program as a baseline. As a result important features like the generation of LHEF when using the NWA are not available. Furthermore, the PDF summation for $\mathcal{KP}$-operator described in Section 4.4 was only implemented in a later version. This becomes especially relevant for the SMEFT application. A particularly large number of parameters is needed in the decomposition of the $\mathcal{KP}$-operator for reweighting even in the SM. A naive extension to the SMEFT case stores a copy of all coefficients for each partial contribution, which massively increases the required storage space. By employing PDF summation one can restrict the use of large LHEF to only truly unique subprocesses, which typically reduces the number of partonic channels by a factor 2 or more. Lastly, the performance optimizations implemented in HELAC-1LOOP cannot be trivially backported into the original framework. It is useful to put all programs on an even footing and thus we implement the UFO interface also into HELAC-DIPOLES using the most up-to-date version as a baseline.

HELAC-DIPOLES by design already separates the evaluation of matrix elements from the internal skeleton construction, which relies on slightly modified modules from the original HELAC-PHEGAS program. A consequence of this structure is that many files are duplicated. For example, the evaluation of the matrix elements is implemented both in `pan1.f90` and `amplitude.f90`. The most straightforward overhaul is thus to first remove all source code that directly originates from HELAC-PHEGAS and then reconstruct the functionality that is actually needed inside HELAC-DIPOLES using stripped down versions of the HELAC-1LOOP source files. Through this procedure all duplicated routines are replaced, performance optimizations regarding the skeleton generation are ported and the two programs are placed on common ground, which simplifies further development. In the next step, a controller was implemented in `helac_controller.f90`, which controls the execution of adopted HELAC-1LOOP routines to generate skeleton files. The current- and vertex-function inclusion is exactly identical to HELAC-1LOOP. The actual evaluation of the matrix elements is still implemented in `amplitude.f90`, which was overhauled using again the respective routines from HELAC-1LOOP. Finally, several HELAC-DIPOLES-specific modules have to be adjusted. The IRC-structure of all SMEFT contributions is identical. Therefore, one can simply promote the weight to an array of weights and keep the remaining structure of the subtraction schemes in place. The book-keeping inside the Monte Carlo integration and the histogram routine are similarly extended. The generation of LHEF has been rewritten into a

dedicated `LHEFhandler.f90` module. The adjustments needed for the reweighting of SMEFT events are similar to the 1-loop case. Individual weights are promoted to an array of weights and coefficients characterizing the logarithmic dependence of the $\mathcal{I}$-operator and $\mathcal{KP}$-operator are now replaced by an array of coefficients. The Wilson coefficients and setting of μ_{EFT} for each specific event alongside the information on the included operators and their mapping to the ordering in the weight array are provided exactly as in the 1-loop case. Helac-Dipoles performs the Monte Carlo integration for positive and negative weights separately, which necessitates the production of two LHEF for each run. As part of this refactoring, we eliminate duplicated code for this purpose and instead control the generated file through a flag. To simplify future development many other improvements have been implemented to improve the overall code quality. The skeleton files now strictly follow the same structure as Helac-1Loop, which enables shared debugging tools between them. Unused routines have been removed throughout many parts of the code and many compiler warnings resolved. Lastly, several features such as the option to integrate up to a specified relative precision and an easier generation of skeleton files have been implemented. Together, these changes lead to improved usability and performance, while also implementing the same UFO interface as outlined for Helac-1Loop.

5.5 Treatment of SMEFT event files in HEPlot

The last necessary piece to enable the full workflow of the Helac-Smeft framework is an update to the reweighting software HEPlot. The additional information provided by the LHEF is read and used to reconstruct the dependence of the partial weights on the scales, Wilson coefficients, PDFs and more. The intermediate Root Ntuple files are expanded based on the number of included SMEFT operators to house the additional information on the matrix elements for each unweighted event. Conceptually the reweighting procedure does not change compared to the previously discussed algorithm in Section 4.4. For each event one unfolds the weight into an array of weights given by

$$w_i = \frac{|\mathcal{M}_i|^2}{\sum_i |\mathcal{M}_i|^2} \, w_{event} \, , \tag{5.11}$$

where the sum over i runs through the combinations of SMEFT Wilson coefficients contained in the given partial weight [17]. By separating these contributions one can implement the running for each Wilson coefficient and recombine it with the corresponding piece. Averaging $f_i \equiv \frac{|\mathcal{M}_i|^2}{\sum_i |\mathcal{M}_i|^2}$ over a large number of events recovers the average contribution of a given partial contribution i to the overall cross section or histogram bin for differential distributions. Combined with the stored normalization factors one can fully reproduce the (differential) cross sections. We note that this statement holds in a statistical sense. Indeed, this approach is less efficient on a per contribution level compared to storing and unweighting just one term, because an additional averaging over the factor f_i has to be performed. When only one weight is present this factor will reduce to 1 in practice. In the future alternative ways of storing this information can be explored. One particular limitation is that we effectively unweight the sum of the SMEFT contribution, which is sensitive to the respective absolute sizes. This can be problematic as the overall weight will be

[17] w_{event} is denoted by XWGTUP in the LHEF standard.

driven only by the ones with the largest weights. Both the phase-space optimization and the unweighting procedure would benefit from a more sophisticated approach. In particular, adopting an approach that is not sensitive to the absolute size of particular contributions would be beneficial when studying the differences of several operators as is done in this thesis. However, we note that while the degradation in statistical accuracy is noticeable, we can still obtain reasonable precision without large CPU time expenses. Beyond the extension of the SM reweighting to SMEFT calculations, we also implement the reweighting of the scale μ_{EFT} at which the Wilson coefficients are defined. To that end, two functionalities are desirable. First, given a Wilson coefficient defined at some high scale such as the new physics scale Λ, dynamical running down to an electroweak scale μ_{EFT} is needed. This allows the evaluation of particular SMEFT correction under some new physics assumptions. Second, we want to evaluate theoretical uncertainties related to the scale setting of μ_{EFT} or more precisely related to precision of the running matrix Γ. To that end, one evaluates the cross-section at multiple central scales of μ_{EFT} and then performs a running to a common scale μ_0. The envelope of these predictions can be used as an estimate of the accuracy of the implemented running matrix. A more detailed description of the procedure is available in Chapter 6. Both concepts are implemented in an automatic way, where we make sure that both aspects can be retroactively inserted based on any event file even if no running was performed during the actual Monte Carlo run. The only ingredient that has to be updated for more general use cases is the running matrix itself. In our current implementation, we insert the running of the $\mathcal{O}_{t\phi}$, $\mathcal{O}_{\phi G}$, $\mathcal{O}_{tG}$, $\mathcal{O}_{tW}$ and $\mathcal{O}_{tB}$ operators by hand up to one-loop precision. However, the current construction can be expanded straightforwardly for different operators. Furthermore, an interface to a multitude of dedicated SMEFT RGE running codes is straightforward and will be implemented in future work.

5.6 Validation and benchmarks

The major overhaul of the framework necessitates thorough validation of all parts of the program. Furthermore, we want to benchmark the performance of the code to determine potential regressions compared to the original state of the program. This section summarizes our efforts to check the correctness of the UFO interface, the updated Helac-Nlo programs and the reweighting pipeline implemented in HEPlot. The validation effort can be summarized into the following points:

- The color-flow decomposition implemented inside the ECRU code via an interface to Form has been checked against an independent implementation that relies on the tensor capabilities of the SymPy python package. The Form based implemented is significantly faster in practice.

- The generated current functions have been tested against the built-in SM implementation in Helac-Nlo on a per-current basis. In addition, we validate the full code generation chain by testing individual phase-space points for several SM processes. Similarly, we compare individual phase-space points using the SMEFT operators $\mathcal{O}_{t\phi}$, $\mathcal{O}_{\phi G}$, $\mathcal{O}_{tG}$ and $\mathcal{O}_{tW}$ against a separate dedicated LO implementation into Helac-Dipoles.

	HELAC-SMEFT [pb]	REF. [95] [pb]	RATIO
$\sigma_{\rm SM}$	$(+5.060 \pm 0.025) \cdot 10^{-1}$	$+5.07 \cdot 10^{-1}$	1.00
$\sigma_{t\phi}$	$(-6.231 \pm 0.030) \cdot 10^{-2}$	$-6.20 \cdot 10^{-2}$	1.01
$\sigma_{\phi G}$	$(+8.732 \pm 0.040) \cdot 10^{-1}$	$+8.72 \cdot 10^{-1}$	1.00
σ_{tG}	$(+5.025 \pm 0.035) \cdot 10^{-1}$	$+5.03 \cdot 10^{-1}$	1.00
$\sigma_{t\phi,\,t\phi}$	$(+1.918 \pm 0.009) \cdot 10^{-3}$	$+1.90 \cdot 10^{-3}$	1.01
$\sigma_{\phi G,\,\phi G}$	$(+1.017 \pm 0.006) \cdot 10^{+0}$	$+1.02 \cdot 10^{+0}$	1.00
$\sigma_{tG,\,tG}$	$(+6.739 \pm 0.036) \cdot 10^{-1}$	$+6.74 \cdot 10^{-1}$	1.00
$\sigma_{t\phi,\,\phi G}$	$(-5.341 \pm 0.024) \cdot 10^{-2}$	$-5.30 \cdot 10^{-2}$	1.01
$\sigma_{t\phi,\,tG}$	$(-3.099 \pm 0.022) \cdot 10^{-2}$	$-3.10 \cdot 10^{-2}$	1.00
$\sigma_{\phi G,\,tG}$	$(+8.496 \pm 0.059) \cdot 10^{-1}$	$+8.59 \cdot 10^{-1}$	0.99

Table 5.3: *Comparison of the integrated cross section results in HELAC-SMEFT framework with Ref. [95] for the $pp \to t\bar{t}H + X$ process at the LHC with $\sqrt{s} =$ 13 TeV. The conventions in Ref. [95] have been adjusted to the SMEFT@NLO convention via suitable rescaling. Taken from Ref. [3].*

- We compare LO and one-loop squared matrix elements for processes with stable top-quarks against GOSAM 3.0 and MADGRAPH5_AMC@NLO for the same selection of SMEFT operators and observe agreement to numerical precision.

- We reproduce results from the $pp \to t\bar{t}H + X$ calculation in Ref. [95]. The center-of-mass energy is set to $\sqrt{s} = 13\,\text{TeV}$. We present a comparison in Table 5.3.

- We explicitly test the independence on the choice of subtraction scheme by comparing cross-sections obtained with the Nagy-Soper and Catani-Seymour schemes implemented in HELAC-DIPOLES. Furthermore, both schemes allow for an additional restriction to the phase-space of the subtraction terms. The dependence on the restriction parameters α_{max} must drop out in the final result, which we confirm numerically.

- The cancellation of IR singularities between the real-emission and virtual contributions is verified for a variety of processes for select phase-space points. In particular, we test both processes involving stable top-quark resonances and decays through the NWA. Full off-shell matrix elements are also considered in the SM [18].

- Leading order and 1-loop color-stripped amplitudes must separately fulfil the QCD Ward identity for the SM and for each SMEFT operator contribution. We investigate several phase-space points explicitly and also use the identity as a measure of numerical stability during the 1-loop event generation. We

[18] The current implementation of SMEFT in the SMEFT@NLO UFO model is not sufficient to support $pp \to t\bar{t}$ within a full off-shell calculation.

note that the amplitude proportional to $C_{\phi G}$ actually fails this test. After careful investigation, we conclude that the finite part of the $ggg H$ counterterm vertex is not correctly implemented in SMEFT@NLO. When recomputing the UV renormalization vertices based on the Lagrangian and the anomalous dimension matrix, we recover gauge invariant results. Details are presented in Section 5.7.

- The robustness of the reweighting procedure is evaluated by preparing several Monte Carlo samples with different scale settings and PDF sets and comparing their results after reweighting to a common scale. In addition, at LO we confirm that histogram files generated using the built-in routines in HELAC-DIPOLES agree with the distributions generated by reweighting LHEF.

While the feature set of the program was expanded for the HELAC-SMEFT framework, it is also important to profile potential regressions to existing functionality. The many optimizations effectively offset any such effect, but it should be noted that the generalizations typically decrease the performance when isolated.

ONE-LOOP SKELETON - CPU TIME (s)			
PROCESS (SM)	HELAC-SMEFT	HELAC-NLO	RATIO
$gg \to gg$	1.25	2.24	0.55
$gg \to ggg$	59.0	284.1	0.21
$gg \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}$	116.8	259.2	0.45
$gg \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H$	1462.4	5198.8	0.28

Table 5.4: *Performance benchmark of the one-loop skeleton generation for various SM processes. Compared are the updated HELAC-SMEFT framework and the HELAC-NLO program. Off-shell matrix elements are considered in all cases and the programs have been compiled using the* `gcc-7.5` *compiler with the* `-O3` *optimization flag enabled. The computer is equipped with an Intel(R) Core(TM) i5-14600 processor.*

The calculation of 1-loop skeleton files is particularly expensive. To get a sense of the state of the program we compare the computation of the skeletons for several SM processes and compare the results in Table 5.4. Many realistic processes are accelerated by a factor 4-5. We note that our optimizations so far have been mostly targeted at the LO recursion. That means that the improvements scale with the complexity of the numerator function, while the construction of loop topologies is often more expensive than before due to the introduced generalizations in the topology generation.

The evaluation of one-loop matrix elements was improved through the introduction of better caching of intermediate currents, while some of the generalizations lead to an increase in computed loop topologies. The latter lead to an increase in one-loop numerators that require reduction with the OPP method. The new setup outperforms the previous version of the code, but further performance improvements can be implemented in the future by more carefully combining topologies when applicable. We compare the performance for individual phase-space points in Table 5.5 and find improvements by 30% to 60% for typical LHC processes with full summation over helicity and color. However, for realistic calculations we always employ

One-Loop Amplitude - CPU Time (s)			
Process (SM)	Helac-Smeft	Helac-Nlo	ratio
$gg \to gg$	2.13	2.42	0.88
$gg \to ggg$	105.4	114.1	0.92
$gg \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}$	5.7	14.4	0.40
$gg \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H$	41.6	61.1	0.68

Table 5.5: *Performance benchmark of the one-loop amplitude calculation for various SM processes. Compared are the updated Helac-Smeft framework and the Helac-Nlo program. Off-shell matrix elements are considered in all cases and the programs have been compiled using the* `gcc-7.5` *compiler with the* `-O3` *optimization flag enabled. Results are fully summed over color and helicity. The computer is equipped with an Intel(R) Core(TM) i5-14600 processor. Taken from Ref. [3].*

One-Loop Amplitude + Stability Check - CPU Time (s)			
Process (SM)	Helac-Smeft	Helac-Nlo	ratio
$gg \to gg$	2.77	4.35	0.63
$gg \to ggg$	134.4	205.4	0.65
$gg \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}$	7.2	22.0	0.33
$gg \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H$	55.7	109.7	0.51

Table 5.6: *Performance benchmark of the one-loop amplitude calculation for various SM processes. Compared are the updated Helac-Smeft framework and the Helac-Nlo program. Off-shell matrix elements are considered in all cases and the programs have been compiled using the* `gcc-7.5` *compiler with the* `-O3` *optimization flag enabled. Results are fully summed over color and helicity. The computer is equipped with an Intel(R) Core(TM) i5-14600 processor.*

the additional checks on the numerical stability of the program for which additional improvements have been implemented. In particular, one can identify which currents are impacted by the replaced wave function in the Ward identity check and only recompute the affected ones. This is especially helpful when the endpoint of the recursion is modified. However, the one-loop reduction still has to be recomputed even in this case. We present the same processes now with the additional stability check enabled in Table 5.6. Matrix elements are fully summed over color and helicity as in Table 5.5. We find that the relative performance improvement improves by about $10 - 20$ percentage points. Lastly, we note that realistic runs will use Monte Carlo techniques to drastically reduce the cost of helicity and color summation. However, the underlying algorithm was not modified. Currently all helicity and color evaluations are independent on each other. As a result, all numbers quoted will scale linearly with the number of evaluated configurations.

5.7 Computation of counterterm Feynman rules

As previously mentioned, we find that the finite part of the $ggg H$ UV counterterm is not correctly implemented in the Smeft@NLO UFO model. Testing our one-loop amplitudes for gauge invariance using the Ward identity, we find a violation in two color-stripped amplitudes. In particular, only the color-flow configurations,

which contain no self-connected gluons are affected. The cancellation of the IR poles between the $\mathcal{I}$-operator and the one-loop amplitude works as expected in our tests, which limits the affected terms to only the finite part. In addition, we find that setting $\mu_R = \mu_{EFT}$ restores gauge-invariance. We therefore expect this term to be proportional to $\log(\mu_R^2/\mu_{EFT}^2)$. Directly editing the counterterms in the SMEFT@NLO UFO model is quite inconvenient. While selecting a unique basis at LO is required in EFT calculations, one has some additional freedom within the counterterms at NLO. All terms must be expressed with respect to the set of Wilson coefficients selected in the LO vertices, but the Feynman rules themselves are not required to be projected on the chosen basis. The SMEFT@NLO model was constructed using an automated tool-chain [105] so additional effort to fully simplify these counterterms was not expended. As a result, some operator combinations might vanish in practice due to EOM or other transformation. To scrutinize the model we thus decide to extract the Feynman rules with a new in-house developed `Python` based tool directly from the Lagrangian. We focus in particular on the five operators $\mathcal{O}_{t\phi}$, $\mathcal{O}_{\phi G}$, $\mathcal{O}_{tG}$, $\mathcal{O}_{tW}$ and $\mathcal{O}_{tB}$, which are finally considered in the analysis of this thesis. We then interface these Feynman rules with our ECRU module to unify the workflow. As a first step, we replicate all LO Feynman rules and find perfect agreement in an analytical comparison. This step also allows us to fully align all sign conventions in the covariant derivatives with the ones used in the UFO model. Next, we recompute the counterterms proportional to $C_{\phi G}$ based on Eq. (3.41). The relevant anomalous-dimension matrix was extracted to first order in α_s in Ref. [117] and can be written as

$$\gamma^{\mathrm{QCD},1} = \frac{1}{3}\begin{pmatrix} -24 & 96\,Y_t & 96\,Y_t^2 & 0 & 0 \\ 0 & -6\,\beta_0^{(6)} & 12\,Y_t & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 8\,g & 8 & 0 \\ 0 & 0 & 8\,g\cos\theta_W - \frac{40}{3}\,g'\sin\theta_W & 0 & 8 \end{pmatrix}, \tag{5.12}$$

where the Wilson coefficients are ordered as $C_i = (C_{t\phi}, C_{\phi G}, C_{tG}, C_{tW}, C_{tZ})$ and the top-quark Yukawa coupling is given by Y_t. We remind the reader that at the one-loop level evanescent operators are not required to obtain finite renormalized results. As previously discussed, one must also modify the SM renormalization to account for the shift of the Higgs fields in the definition of $\mathcal{O}_{t\phi}$ and $\mathcal{O}_{\phi G}$. The relevant terms are given in Ref. [95]. In practice, all additional terms are proportional to C_{tG}. Because we are particularly interested in the $C_{\phi G}$ terms, we can thus ignore them. If we wanted to fully include these terms, we would have to replicate the SM Lagrangian as well. As a consistency check we find that the $C_{t\phi}$ counterterms extracted using our module are identical to those found in the SMEFT@NLO UFO model. When replacing the $C_{\phi G}$ UV counterterms with our derived expression we find that gauge invariance is restored. For our final setup we thus combine the $C_{\phi G}$ rational terms of second kind from the UFO model with the counterterms that we extract using the Lagrangian. The other operators are taken as is from the UFO model as well.

Chapter 6

The $pp \to t\bar{t}H$ process in SMEFT

In this chapter, we discuss the calculation of the $pp \to t\bar{t}H + X$ process in the dileptonic decay channel using the HELAC-SMEFT framework. Additionally, we repeat the same calculation with stable top-quarks to better understand the differences of the two top-quark modelling approaches. We first outline the computational setup of both calculations and identify the most important SMEFT operators for this process. We then describe the results across several sections. First, we focus on the integrated fiducial cross sections within both approaches. Next, we study differential distributions in the NWA and investigate differences and similarities of kinematical features between the SM and several of the SMEFT contributions. We also briefly compare the impact of fixed and dynamical scale choices on differential cross section predictions. Finally, we carefully investigate the impact of the two top-quark modelling prescriptions on differential distributions and highlight again where the SM and SMEFT contributions behave differently.

6.1 Computational setup

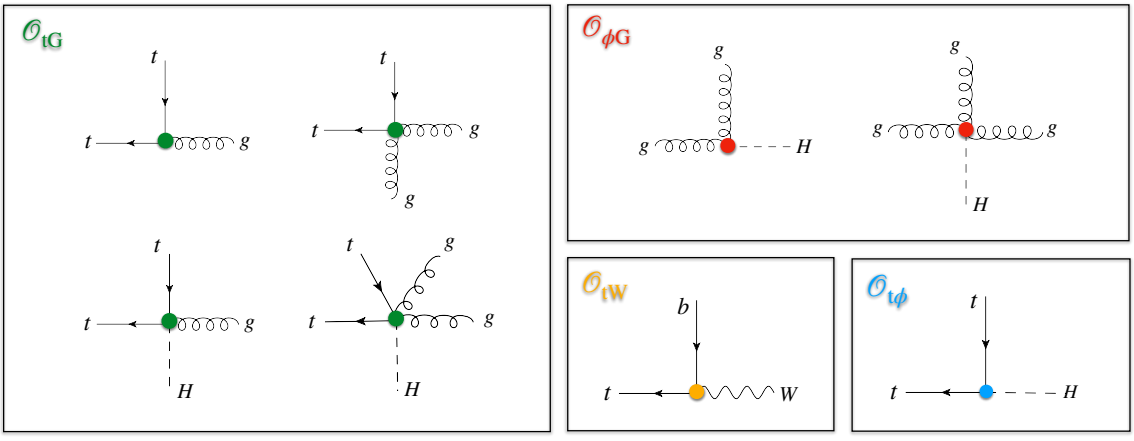


Figure 6.1: *SMEFT vertices entering the $pp \to t\bar{t}H + X \to W^+W^-b\bar{b}H + X \to e^+\nu_e\mu^-\bar{\nu}_\mu b\bar{b}H + X$ process at NLO QCD. The included operators are $\mathcal{O}_{tG}$, $\mathcal{O}_{\phi G}$, $\mathcal{O}_{t\phi}$ and $\mathcal{O}_{tW}$. Taken from Ref. [3].*

We study the impact of realistic final-state selections and top-quark decays on SMEFT predictions using the $pp \to t\bar{t}H$ production process. To that end, we

consider the

$$pp \to t\bar{t}H + X\,, \tag{6.1}$$

production process, where we treat the top-quarks as stable particles and the

$$pp \to t\bar{t}H + X \to W^+W^-b\bar{b}H + X \to e^+\nu_e\mu^-\bar{\nu}_\mu b\bar{b}H + X \tag{6.2}$$

process, where we treat the top quark and W resonances using the NWA. We refer to the former calculation as STABLE and the latter as NWA throughout this chapter. The first approach allows for the study of production rates, but does not include predictions about the impact of the top-quark decays and fiducial phase-space cuts. All results are computed at NLO QCD, meaning up to $\mathcal{O}(\alpha_s^3\alpha)$ and $\mathcal{O}(\alpha_s^3\alpha^5)$ respectively, and SMEFT contributions are included consistently in production and decay in the NWA computation. We present results for one specific assignment of lepton flavors in the decay chain, but the results for arbitrary configurations of electron and muon flavors $\ell = e^\pm$, $\mu^\pm$ can be obtained by rescaling all results by a factor of 4. The center-of-mass energy is set to the Run III value of $\sqrt{s} = 13.6\,\text{TeV}$.

We briefly recap important effective SMEFT operators that contribute to this process. The $\mathcal{O}_G$ modifies gluon self-interactions, which are present in the production process of any $t\bar{t}$-associated process. However, it can be well constrained using multi-jet production as well as the $t\bar{t}$ process at the LHC. To simplify our study, we assume that these measurements sufficiently constrain such contributions and we can therefore neglect them. The dominant operators in the production process are given by $\mathcal{O}_{t\phi}$, $\mathcal{O}_{\phi G}$ and $\mathcal{O}_{tG}$, which are all considered in this thesis. Top-quark decays receive corrections from $\mathcal{O}_{tG}$ at NLO in QCD through QCD loops and real-emission contributions. In addition, the $\mathcal{O}_{tW}$ operator provides a direct tree-level modification of the chirality structure of the tWb vertex. Other contributions to the top-quark decay are given by the $\mathcal{O}_{\phi Q}^{(1)}$ and $\mathcal{O}_{\phi Q}^{(3)}$ operators, which we do not consider here, as their underlying spinor structure is the same as the SM. Many other operators can also be included. In particular, four-fermion operators are relevant to the process. However, their contributions are disconnected from the Higgs boson. The measurement of their respective Wilson coefficients is better performed using $pp \to t\bar{t}$, which benefits from a much larger cross section. Closure of the renormalization procedure would generally require the inclusion of the $\mathcal{O}_{tB}$ operator as well. However, we will be working in the NWA where it can not enter at LO. Since the counterterms are also proportional to the LO amplitude we can thus neglect it fully. Finally, we end up with the following four operators for this study:

$$\begin{aligned}
\mathcal{O}_{t\phi} &= \left(\phi^\dagger\phi - \tfrac{v^2}{2}\right)\bar{Q}\,t\,\tilde{\phi} + h.c.\,,\\
\mathcal{O}_{\phi G} &= \left(\phi^\dagger\phi - \tfrac{v^2}{2}\right)G^{a\mu\nu}\,G^a_{\mu\nu}\,,\\
\mathcal{O}_{tG} &= ig_s\left(\bar{Q}\sigma^{\mu\nu}t^a t\right)\tilde{\phi}\,G^a_{\mu\nu} + h.c.\\
\mathcal{O}_{tW} &= i\left(\bar{Q}\sigma^{\mu\nu}\sigma^I t\right)\tilde{\phi}\,W^I_{\mu\nu} + h.c.\,.
\end{aligned} \tag{6.3}$$

Examples of the vertices induced by these operators relevant to the $pp \to t\bar{t}H$ process are given in Figure 6.1. Furthermore, Feynman diagrams of the $pp \to t\bar{t}H + X \to W^+W^-b\bar{b}H + X$ process are shown at LO in Figure 6.2, where we have omitted the leptonic decays of the W bosons for brevity. These diagrams contain SMEFT corrections to both the production and decay of the process. The operator

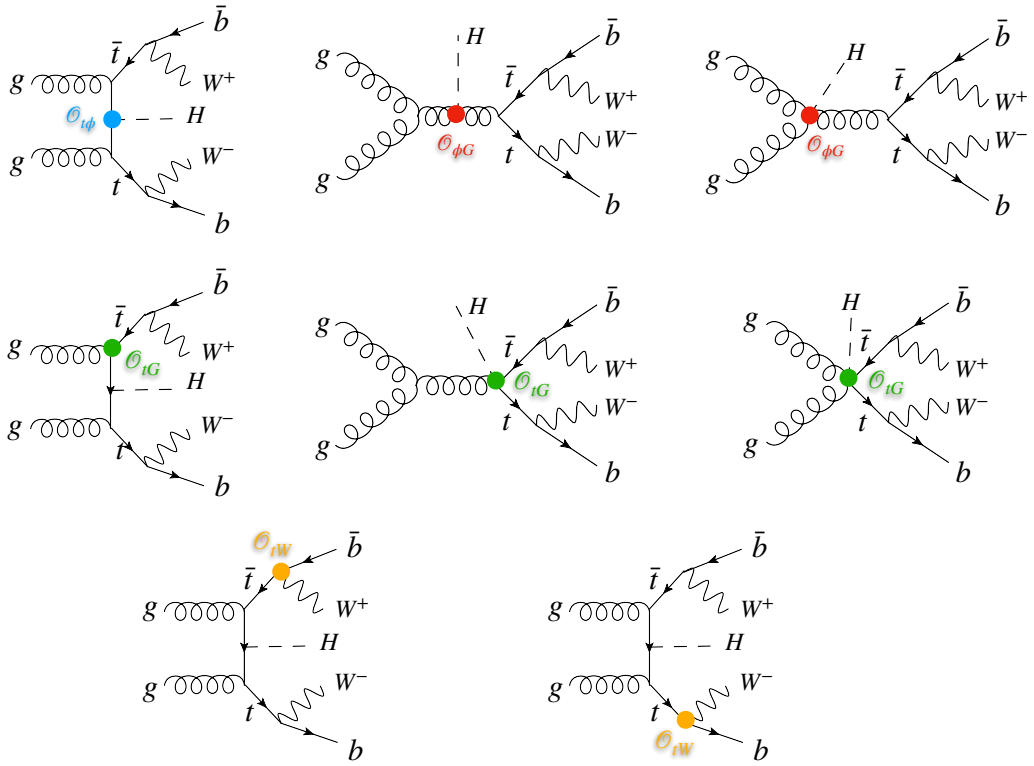


Figure 6.2: *Representative LO diagrams of the SMEFT contributions to the* $pp \to t\bar{t}H + X \to W^+W^-b\bar{b}H + X$ *process. The subsequent* W *boson decays are not displayed in the figure to increase the clarity. The operators* $\mathcal{O}_{tG}$, $\mathcal{O}_{\phi G}$, $\mathcal{O}_{t\phi}$ *and* $\mathcal{O}_{tW}$ *are included and the amplitude is truncated at* $\mathcal{O}(1/\Lambda^2)$. *Taken from Ref. [3].*

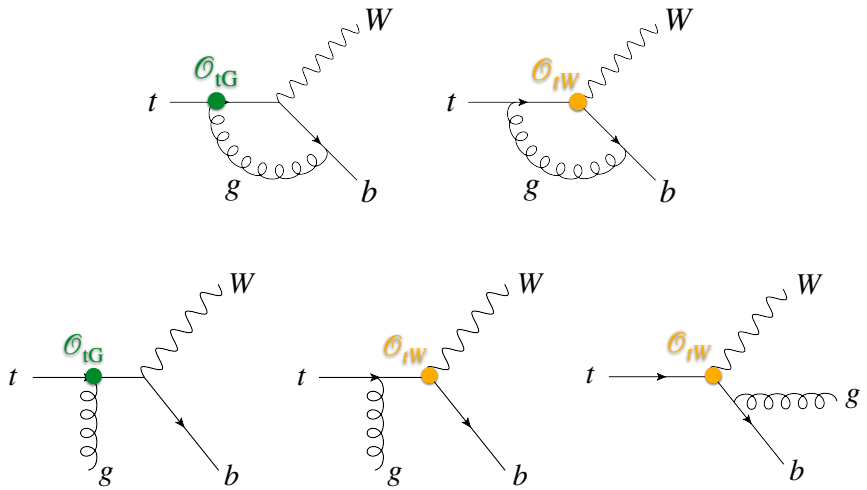


Figure 6.3: *Examples of* $\mathcal{O}_{tG}$ *and* $\mathcal{O}_{tW}$ *vertices entering the top-quark decays at NLO QCD in SMEFT. The* W *boson decays are omitted to increase the clarity of the figure.*

insertions into the top-quark decays at NLO are shown in Figure 6.3, where we include both $\mathcal{O}_{tG}$ and $\mathcal{O}_{tW}$, respectively. To demonstrate the complexity of the calculation we present the number of Feynman diagrams involved at LO and NLO in Table 6.1 for both the calculation with stable top quarks (right) and within in the NWA (left). We note that the HELAC-SMEFT framework does not compute results using these Feynman diagrams and instead relies on recursive methods. However, they can still be useful to understand the differences between the operators and their impact on the required computational resources. The diagrams are organized according to the SMEFT Wilson coefficients included in the respective amplitude. The SM is also shown for comparison purposes. The real-emission channels are further split into diagrams in which the extra-parton is emitted from the production process or the top-quark decays.

	SM	$C_{t\phi}$	$C_{\phi G}$	C_{tG}	C_{tW}
$gg \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}H$					
Born	8	8	8	22	16
one-loop	178	178	289	774	356
$q\bar{q}/\bar{q}q \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}H$					
Born	2	2	1	3	4
one-loop	43	43	42	122	86
$gg \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}Hg$					
Real prod.	50	50	74	183	100
Real decay	16	16	16	52	32
$q\bar{q}/\bar{q}q \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}Hg$					
Real prod.	12	12	12	28	24
Real decay	4	4	2	8	8
$qg/gq \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}Hq$					
Real prod.	12	12	12	28	24

	SM	$C_{t\phi}$	$C_{\phi G}$	C_{tG}
$gg \to t\bar{t}H$				
Born	8	8	8	22
one-loop	162	162	273	714
$q\bar{q}/\bar{q}q \to t\bar{t}H$				
Born	2	2	1	3
one-loop	39	39	40	112
$gg \to t\bar{t}Hg$				
Real emission	50	50	74	183
$q\bar{q}/\bar{q}q \to t\bar{t}Hg$				
Real emission	12	12	12	28
$qg/gq \to t\bar{t}Hq$				
Real emission	12	12	12	28

Table 6.1: *Comparison of number of Feynman diagrams for the NLO calculation of $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process in the NWA (left) and $pp \to t\bar{t}H + X$ process (right). The contributions are split according to the Wilson-coefficient content. Born, one-loop and real-radiation contributions are shown separately. The real-radiation diagrams are further split into terms where the extra radiation is emitted in the production and the decay matrix elements. Taken from Ref. [3].*

We note that this set of operators does not modify the SM relations between the gauge couplings, the Higgs vacuum expectation value or the definition of the weak mixing angle. As previously discussed this is generally not true in SMEFT. This allows us to use SM relation to derive the internal constants from the input parameters in the context of this study. We derive the SM parameters in the G_μ-

scheme using the SM input values

$$\begin{aligned} G_\mu &= 1.166378 \cdot 10^{-5}\ \text{GeV}^{-2}\,, & m_t &= 172.5\ \text{GeV}\,, \\ m_W &= 80.379\ \text{GeV}\,, & \Gamma_W &= 2.0972\ \text{GeV}\,, \\ m_Z &= 91.1876\ \text{GeV}\,, & m_H &= 125.0\ \text{GeV}\,. \end{aligned} \tag{6.4}$$

The values of the electroweak coupling α and the weak-mixing angle $\sin\theta_W$ are derived from the input constants. In accordance with the conventions of SMEFT@NLO, we work in the five-flavor scheme, i.e. we keep the bottom quark massless, and use a diagonal CKM matrix. The SM LO and NLO top-quark width are derived from the input parameters based on the equations in Ref. [230]. In this study, we keep the top-quark width fixed to the corresponding SM value. Numerically, we use

$$\Gamma_t^{\text{LO}} = 1.4806842\,\text{GeV}\,, \qquad \Gamma_t^{\text{NLO}} = 1.3535983\,\text{GeV}\,. \tag{6.5}$$

The numerical results presented in this thesis include SMEFT effects in the matrix elements, while the Wilson-coefficient dependence of the total top-quark width is kept separate. In the NWA, this dependence enters through the factorized inverse-width factors associated with the top-quark resonances and can be incorporated by a consistent re-expansion at the level of the binned distributions. To that end, one decomposes the top-quark width into a SM part and SMEFT corrections,

$$\Gamma_t^{\text{SMEFT}} = \Gamma_t^{\text{SM}} + \delta\Gamma_t^{(6)}\,, \tag{6.6}$$

where $\delta\Gamma_t^{(6)}$ is linear in the dimension-6 Wilson coefficients. In the NWA, each resonance gives rise to one factor of the inverse decay width, which can be expanded as

$$\frac{1}{\Gamma_t^{\text{SMEFT}}} = \frac{1}{\Gamma_t^{\text{SM}} + \delta\Gamma_t^{(6)}} = \frac{1}{\Gamma_t^{\text{SM}}}\left(1 - \frac{\delta\Gamma_t^{(6)}}{\Gamma_t^{\text{SM}}} + \left(\frac{\delta\Gamma_t^{(6)}}{\Gamma_t^{\text{SM}}}\right)^2 + \dots\right). \tag{6.7}$$

The additional terms would then be obtained by combining this expansion with the cross-section contributions and truncating at the desired order in $1/\Lambda$. The NLO QCD top-quark width in SMEFT has been reported in Ref. [126]. The C_{tG} contribution to the decay width is loop-induced and numerically small, whereas the C_{tW} terms can be more important and should be accounted for in experimental studies. This treatment relies on the factorized resonance structure of the NWA and does not directly apply to calculations with non-factorizable corrections. We leave the treatment of SMEFT corrections to the top-quark width to future work.

We define a fiducial phase-space region for the NWA calculation by requiring two charged leptons and at least two b-jets. Jet observables are computed by clustering the final state partons with rapidity $|y| < 5$ with the anti-k_T jet algorithm [140], where we set the jet radius parameter $R = 0.4$. The kinematics of the final state jets is restricted by

$$\begin{aligned} p_{T,b} &> 25\,\text{GeV}\,, & |y_b| &< 2.5\,, & \Delta R_{bb} &> 0.4\,, \\ p_{T,\ell} &> 25\,\text{GeV}\,, & |y_\ell| &< 2.5\,, & \Delta R_{\ell\ell} &> 0.4\,, \end{aligned} \tag{6.8}$$

where ℓ can be e or μ and $\Delta R_{ij} = \sqrt{(\Delta y_{ij})^2 + (\Delta\phi_{ij})^2}$ is the distance in the rapidity-transverse azimuthal angle plane. We employ the LO and NLO NNPDF3.1 PDF

set [137], which are interfaced using LHAPDF6 [224] program. The latter also provides the running of $\alpha_s(\mu)$, where we select $\alpha_s(m_Z) = 0.118$ at both LO and NLO.

We organize the calculation into SM as well as linear and quadratic partial SMEFT contributions, which results in the following decomposition:

$$\begin{aligned}\sigma(\mu_R, \mu_F, \mu_{EFT}) = \sigma_{\text{SM}}(\mu_R, \mu_F) + \sum_i \Sigma_i^{\text{SMEFT}}(\mu_R, \mu_F, \mu_{EFT}) \\ + \sum_{i,j} \Sigma_{ij}^{\text{SMEFT}}(\mu_R, \mu_F, \mu_{EFT})\,,\end{aligned} \tag{6.9}$$

where we made the scale μ_{EFT}, at which the operators are defined, explicit. We note that the dependence on μ_{EFT} does not fully vanish in Eq. (6.9), because the evolution of the Wilson coefficients is not known to all orders. The mixing matrix for this set of operators was given in Eq. (5.12) to first order in α_s, where the last entry corresponds to the C_{tB} coefficient that does not enter our study. The corresponding running matrix Γ is computed analytically using Eq. (3.45), where we set the number of active flavors to $n_F = 5$. We use this definition of Γ to perform the running between two values of the μ_{EFT} scale.

We compare two choices for the scale setting. In both cases we consider a fixed setting $\mu_{EFT} = m_t$ for the value at which the Wilson coefficients are defined. This ensures that they are phase-space point independent and allow us to factor them out universally, while retaining comparability between the two setups. For the renormalization and factorization scales, μ_R and μ_F, we study and compare two setups. First a fixed scale setup given by

$$\mu_0 = \mu_R = \mu_F = m_t + \frac{m_H}{2}\,, \qquad \mu_0 = \mu_{EFT} = m_t\,, \tag{6.10}$$

as well as dynamical scale setup centered around

$$\mu_0 = \mu_R = \mu_F = \frac{H_T}{2}\,, \qquad \mu_0 = \mu_{EFT} = m_t\,, \tag{6.11}$$

where the H_T is the scalar sum of the transverse momenta in the final state defined according to

$$\begin{aligned} H_T^{\text{NWA}} &= p_T(e^+) + p_T(\mu^-) + p_T(b_1) + p_T(b_2) + p_T^{miss} + p_T(H)\,, \\ H_T^{\text{STABLE}} &= p_T(t) + p_T(\bar{t}) + p_T(H)\,, \end{aligned} \tag{6.12}$$

for the calculations in the NWA and with stable top quarks, respectively. For a fixed choice of μ_{EFT}, we can further parametrize the results following the approach used in Ref. [95] and obtain

$$\begin{aligned} \Sigma_i^{\text{SMEFT}}(\mu_R, \mu_F, \mu_{EFT}) &= \frac{1\,\text{TeV}^2}{\Lambda^2}\, C_i(\mu_{EFT})\, \sigma_i(\mu_R, \mu_F, \mu_{EFT})\,, \\ \Sigma_{ij}^{\text{SMEFT}}(\mu_R, \mu_F, \mu_{EFT}) &= \frac{1\,\text{TeV}^4}{\Lambda^4}\, C_i(\mu_{EFT})\, C_j(\mu_{EFT})\, \sigma_{ij}(\mu_R, \mu_F, \mu_{EFT})\,, \end{aligned} \tag{6.13}$$

where we have additionally factored out powers of $1\,\text{TeV}/\Lambda$. This parametrization allows for trivial changes in the new physics scale Λ, but we note that the partial cross sections σ_i and σ_{ij} effectively include powers of $(1\text{TeV})^{-2}$ and $(1\text{TeV})^{-4}$, respectively.

We estimate theoretical uncertainties due to the truncation of the α_s expansion of the amplitude using 7-point scale variation, that is, we evaluate observables at

$$\left(\frac{\mu_R}{\mu_0}, \frac{\mu_F}{\mu_0}\right) = \Big\{ (2,1), (0.5,1), (1,2), (1,1), (1,0.5), (2,2), (0.5,0.5) \Big\}, \tag{6.14}$$

and compute the envelope. Missing higher order contributions to the RGE evolution and mixing of the operators can be estimated by varying the μ_{EFT} scale. In particular, given some central value $\mu_{EFT} = \mu_0$ and a different scale $\mu_{EFT} = \mu$, we can write the cross sections as

$$\begin{aligned} \sigma(\mu_0) &= \sigma_{\text{SM}} + \sum_i \frac{1\,\text{TeV}^2}{\Lambda^2} C_i(\mu_0)\, \sigma_i(\mu_0) + \sum_{i,j} \frac{1\,\text{TeV}^4}{\Lambda^4} C_i(\mu_0)\, C_j(\mu_0)\, \sigma_{ij}(\mu_0)\,, \\ \sigma(\mu) &= \sigma_{\text{SM}} + \sum_i \frac{1\,\text{TeV}^2}{\Lambda^2} C_i(\mu)\, \sigma_i(\mu) + \sum_{i,j} \frac{1\,\text{TeV}^4}{\Lambda^4} C_i(\mu)\, C_j(\mu)\, \sigma_{ij}(\mu)\,, \end{aligned} \tag{6.15}$$

which can be put on a common ground by running the Wilson coefficients $C_i(\mu)$ back to the scale μ_0. Inserting

$$C_i(\mu) = \Gamma_{ij}(\mu, \mu_0)\, C_j(\mu_0) \tag{6.16}$$

into $\sigma(\mu)$ in Eq. (6.15) yields

$$\sigma(\mu) = \sigma_{\text{SM}} + \sum_i \frac{1\,\text{TeV}^2}{\Lambda^2} C_i(\mu_0)\, \sigma_i(\mu_0, \mu) + \sum_{ij} \frac{1\,\text{TeV}^4}{\Lambda^4} C_i(\mu_0)\, C_j(\mu_0)\, \sigma_{ij}(\mu_0, \mu)\,, \tag{6.17}$$

where we define $\sigma_i(\mu_0, \mu)$ and $\sigma_{ij}(\mu_0, \mu)$ in terms of the running matrix Γ_{ij} as

$$\begin{aligned} \sigma_i(\mu_0, \mu) &= \Gamma_{ji}(\mu, \mu_0)\, \sigma_j(\mu)\,, \\ \sigma_{ij}(\mu_0, \mu) &= \Gamma_{ki}(\mu, \mu_0)\, \Gamma_{lj}(\mu, \mu_0)\, \sigma_{kl}(\mu)\,. \end{aligned} \tag{6.18}$$

This setup transfers the RGE running from the Wilson coefficients back into the cross section and allows us to factor out common coefficients. We can then directly study the uncertainties using σ_i and σ_{ij}. To that end, we evaluate Eq. (6.18) at three values of μ_{EFT} given by

$$\mu_{EFT} = 2\mu_0\,, \qquad \mu_{EFT} = \mu_0\,, \qquad \mu_{EFT} = 0.5\mu_0\,, \tag{6.19}$$

and quote the envelope of three results as the theoretical uncertainty related to the precision of the RGE running in the calculation. Theoretical uncertainties related to the choice of PDF set and the internal uncertainties of individual sets are independent of the SMEFT operators in the amplitude. Ref. [95] provides estimate for PDF uncertainties for the $pp \to t\bar{t}H$ process in SMEFT with stable top quarks. The SM $pp \to t\bar{t}H$ including top-quark decays has been studied in Ref. [74], where estimates for the theoretical uncertainties were provided. We do not recalculate them here, but expect them to be negligible compared to scale uncertainties at NLO QCD. Calculations with higher precision in the perturbative expansion would need include them more carefully. We note, however, that our reweighting setup is already fully equipped to estimate these uncertainties without any additional Monte Carlo runs.

6.2 Integrated results

This section is dedicated to the study of phenomenological differences of the two top-quark modelling approaches considered in this thesis. The stable top-quark calculation, namely $pp \to t\bar{t}H + X$, and the NWA calculation in the dileptonic decay channel, i.e. $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$, are compared on the integrated level. In particular, we focus on effects induced by the fiducial phase-space cuts in the NWA as well as the genuine SMEFT corrections to the top-quark decays.

$pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ (NWA)					
	σ^{LO}[fb]	$\sigma^{\text{LO}}/\sigma^{\text{LO}}_{\text{SM}}$	σ^{NLO}[fb]	$\sigma^{\text{NLO}}/\sigma^{\text{NLO}}_{\text{SM}}$	$\mathcal{K}$
σ_{SM}	$2.3166^{+0.7068\,+0.0000}_{-0.5044\,-0.0000}$	$1.000^{+0.000\,+0.000}_{-0.000\,-0.000}$	$2.7680^{+0.0560\,+0.0000}_{-0.1898\,-0.0000}$	$1.000^{+0.000\,+0.000}_{-0.000\,-0.000}$	1.19
$\sigma_{t\phi}$	$-0.2835^{+0.0617\,+0.0126}_{-0.0865\,-0.0145}$	$-0.122^{+0.000\,+0.005}_{-0.000\,-0.006}$	$-0.3519^{+0.0270\,+0.0037}_{-0.0091\,-0.0050}$	$-0.127^{+0.001\,+0.001}_{-0.002\,-0.002}$	1.24
$\sigma_{\phi G}$	$1.6262^{+0.5137\,+0.2089}_{-0.3638\,-0.1730}$	$0.702^{+0.006\,+0.090}_{-0.005\,-0.075}$	$2.5292^{+0.2865\,+0.1154}_{-0.3098\,-0.0954}$	$0.914^{+0.094\,+0.042}_{-0.053\,-0.034}$	1.56
σ_{tG}	$2.4217^{+0.7568\,+0.0108}_{-0.5374\,-0.0137}$	$1.045^{+0.006\,+0.005}_{-0.006\,-0.006}$	$2.8249^{+0.0460\,+0.0133}_{-0.1774\,-0.0157}$	$1.021^{+0.006\,+0.005}_{-0.016\,-0.006}$	1.17
σ_{tW}	$0.7238^{+0.2214\,+0.0111}_{-0.1579\,-0.0120}$	$0.312^{+0.000\,+0.005}_{-0.000\,-0.005}$	$0.8719^{+0.0182\,+0.0025}_{-0.0615\,-0.0041}$	$0.315^{+0.001\,+0.001}_{-0.001\,-0.001}$	1.20
$\sigma_{t\phi,t\phi}$	$0.0087^{+0.0026\,+0.0009}_{-0.0019\,-0.0008}$	$0.004^{+0.000\,+0.000}_{-0.000\,-0.000}$	$0.0112^{+0.0005\,+0.0003}_{-0.0009\,-0.0003}$	$0.004^{+0.000\,+0.000}_{-0.000\,-0.000}$	1.29
$\sigma_{t\phi,\phi G}$	$-0.0995^{+0.0223\,+0.0146}_{-0.0314\,-0.0185}$	$-0.043^{+0.000\,+0.006}_{-0.000\,-0.008}$	$-0.1592^{+0.0203\,+0.0095}_{-0.0196\,-0.0111}$	$-0.058^{+0.004\,+0.003}_{-0.006\,-0.004}$	1.60
$\sigma_{t\phi,tG}$	$-0.1482^{+0.0329\,+0.0060}_{-0.0463\,-0.0067}$	$-0.064^{+0.000\,+0.003}_{-0.000\,-0.003}$	$-0.1794^{+0.0128\,+0.0009}_{-0.0034\,-0.0014}$	$-0.065^{+0.000\,+0.000}_{-0.000\,-0.000}$	1.21
$\sigma_{t\phi,tW}$	$-0.0443^{+0.0097\,+0.0013}_{-0.0135\,-0.0015}$	$-0.019^{+0.000\,+0.001}_{-0.000\,-0.001}$	$-0.0553^{+0.0043\,+0.0004}_{-0.0016\,-0.0005}$	$-0.020^{+0.000\,+0.000}_{-0.000\,-0.000}$	1.25
$\sigma_{\phi G,\phi G}$	$0.8178^{+0.2987\,+0.1799}_{-0.2032\,-0.1361}$	$0.353^{+0.016\,+0.078}_{-0.014\,-0.059}$	$1.5299^{+0.2443\,+0.1528}_{-0.2377\,-0.1277}$	$0.553^{+0.082\,+0.055}_{-0.052\,-0.046}$	1.87
$\sigma_{\phi G,tG}$	$1.5954^{+0.5561\,+0.1352}_{-0.3835\,-0.1175}$	$0.689^{+0.023\,+0.058}_{-0.020\,-0.051}$	$2.4966^{+0.2465\,+0.0653}_{-0.3088\,-0.0533}$	$0.902^{+0.080\,+0.024}_{-0.053\,-0.019}$	1.56
$\sigma_{\phi G,tW}$	$0.2548^{+0.0807\,+0.0280}_{-0.0571\,-0.0236}$	$0.110^{+0.001\,+0.012}_{-0.001\,-0.010}$	$0.3980^{+0.0458\,+0.0155}_{-0.0491\,-0.0126}$	$0.144^{+0.015\,+0.006}_{-0.008\,-0.005}$	1.56
$\sigma_{tG,tG}$	$3.4828^{+1.2967\,+0.0660}_{-0.8776\,-0.0730}$	$1.503^{+0.077\,+0.028}_{-0.066\,-0.032}$	$4.3464^{+0.0425\,+0.0262}_{-0.3089\,-0.0365}$	$1.570^{+0.007\,+0.009}_{-0.065\,-0.013}$	1.25
$\sigma_{tG,tW}$	$0.3795^{+0.1189\,+0.0079}_{-0.0843\,-0.0088}$	$0.164^{+0.001\,+0.003}_{-0.001\,-0.004}$	$0.4448^{+0.0074\,+0.0033}_{-0.0285\,-0.0045}$	$0.161^{+0.001\,+0.001}_{-0.002\,-0.002}$	1.17
$\sigma_{tW,tW}$	$0.0754^{+0.0231\,+0.0023}_{-0.0165\,-0.0025}$	$0.033^{+0.000\,+0.001}_{-0.000\,-0.001}$	$0.0906^{+0.0019\,+0.0005}_{-0.0063\,-0.0009}$	$0.033^{+0.000\,+0.000}_{-0.000\,-0.000}$	1.20

Table 6.2: *Integrated cross sections at LO and NLO for the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process with center-of-mass energy $\sqrt{s} = 13.6$ TeV. The SMEFT contributions are split into linear, quadratic and interference terms. The scales are set to $\mu_R = \mu_F = m_t + m_H/2$ as well as $\mu_{EFT} = m_t$ and the NNPDF3.1 PDF set is employed. The first uncertainty band corresponds to the 7-point scale variation of μ_R and μ_F, while the second one estimates the dependence on μ_{EFT}. Taken from Ref. [3].*

We present the integrated cross section at LO and NLO accuracy with the fixed scale setup $\mu_R = \mu_F = m_t + m_H/2$ in Tables 6.2 and 6.3 for calculations within the NWA and with stable top-quarks, respectively. The corresponding results for the dynamical scale choice $\mu_R = \mu_F = H_T/2$ are given in Tables 6.4 and 6.5. In all cases we define the Wilson coefficients at $\mu_{EFT} = m_t$ and employ the NNPDF3.1 PDF set. We organize the contributions according to the contained Wilson coefficients by utilising Eq. (6.9) and (6.13). The SM results is provided alongside the linear terms σ_i, the interference contribution σ_{ij} as well as the quadratic contribution σ_{ii}, where the factor 2 originating from $i \leftrightarrow j$ has already been included. In addition, we explicitly provide the prediction to the SM, where uncertainties have been consistently evaluated assuming correlated scales. The first super and subscript of each entry correspond to the scale uncertainties obtained by standard 7-point variation

$pp \to t\bar{t}H + X$ (STABLE)					
	σ^{LO}[pb]	$\sigma^{\text{LO}}/\sigma^{\text{LO}}_{\text{SM}}$	σ^{NLO}[pb]	$\sigma^{\text{NLO}}/\sigma^{\text{NLO}}_{\text{SM}}$	$\mathcal{K}$
σ_{SM}	$0.4424^{+0.1337\,+0.0000}_{-0.0956\,-0.0000}$	$1.000^{+0.000\,+0.000}_{-0.000\,-0.000}$	$0.5623^{+0.0327\,+0.0000}_{-0.0519\,-0.0000}$	$1.000^{+0.000\,+0.000}_{-0.000\,-0.000}$	1.27
$\sigma_{t\phi}$	$-0.0541^{+0.0117\,+0.0024}_{-0.0164\,-0.0028}$	$-0.122^{+0.000\,+0.005}_{-0.000\,-0.006}$	$-0.0713^{+0.0071\,+0.0009}_{-0.0051\,-0.0012}$	$-0.127^{+0.001\,+0.002}_{-0.002\,-0.002}$	1.32
$\sigma_{\phi G}$	$0.3023^{+0.0945\,+0.0392}_{-0.0671\,-0.0324}$	$0.683^{+0.005\,+0.088}_{-0.005\,-0.073}$	$0.4941^{+0.0724\,+0.0247}_{-0.0683\,-0.0204}$	$0.879^{+0.073\,+0.044}_{-0.044\,-0.036}$	1.63
σ_{tG}	$0.4501^{+0.1393\,+0.0004}_{-0.0992\,-0.0008}$	$1.017^{+0.006\,+0.001}_{-0.005\,-0.002}$	$0.5641^{+0.0268\,+0.0022}_{-0.0500\,-0.0024}$	$1.003^{+0.004\,+0.004}_{-0.010\,-0.004}$	1.25
$\sigma_{t\phi,t\phi}$	$0.0017^{+0.0005\,+0.0002}_{-0.0004\,-0.0001}$	$0.004^{+0.000\,+0.000}_{-0.000\,-0.000}$	$0.0023^{+0.0002\,+0.0001}_{-0.0002\,-0.0001}$	$0.004^{+0.000\,+0.000}_{-0.000\,-0.000}$	1.36
$\sigma_{t\phi,\phi G}$	$-0.0185^{+0.0041\,+0.0027}_{-0.0058\,-0.0035}$	$-0.042^{+0.000\,+0.006}_{-0.000\,-0.008}$	$-0.0311^{+0.0044\,+0.0020}_{-0.0048\,-0.0023}$	$-0.055^{+0.003\,+0.004}_{-0.005\,-0.004}$	1.68
$\sigma_{t\phi,tG}$	$-0.0275^{+0.0061\,+0.0012}_{-0.0085\,-0.0014}$	$-0.062^{+0.000\,+0.003}_{-0.000\,-0.003}$	$-0.0357^{+0.0034\,+0.0003}_{-0.0022\,-0.0004}$	$-0.064^{+0.000\,+0.001}_{-0.000\,-0.001}$	1.30
$\sigma_{\phi G,\phi G}$	$0.1506^{+0.0548\,+0.0332}_{-0.0373\,-0.0251}$	$0.340^{+0.016\,+0.075}_{-0.014\,-0.057}$	$0.2991^{+0.0581\,+0.0318}_{-0.0511\,-0.0263}$	$0.532^{+0.068\,+0.057}_{-0.046\,-0.047}$	1.99
$\sigma_{\phi G,tG}$	$0.2958^{+0.1027\,+0.0257}_{-0.0709\,-0.0223}$	$0.669^{+0.023\,+0.058}_{-0.020\,-0.050}$	$0.4965^{+0.0704\,+0.0151}_{-0.0712\,-0.0125}$	$0.883^{+0.070\,+0.027}_{-0.050\,-0.022}$	1.68
$\sigma_{tG,tG}$	$0.5852^{+0.2145\,+0.0106}_{-0.1459\,-0.0118}$	$1.323^{+0.065\,+0.024}_{-0.056\,-0.027}$	$0.7798^{+0.0337\,+0.0056}_{-0.0749\,-0.0074}$	$1.387^{+0.005\,+0.010}_{-0.040\,-0.013}$	1.33

Table 6.3: *Integrated cross sections at LO and NLO for the $pp \to t\bar{t}H + X$ process with center-of-mass energy $\sqrt{s} = 13.6$ TeV. The SMEFT contributions are split into linear, quadratic and interference terms. The scales are set to $\mu_R = \mu_F = m_t + m_H/2$ as well as $\mu_{EFT} = m_t$ and the NNPDF3.1 PDF set is employed. The first uncertainty band corresponds to the 7-point scale variation of μ_R and μ_F, while the second one estimates the dependence on μ_{EFT}. Taken from Ref. [3].*

of μ_R and μ_F. The second pair is the envelope obtained through the variation of μ_{EFT}. Lastly, the integrated $\mathcal{K}$-factors, defined as the ratio of NLO corrections to the LO predictions, are shown in the last column.

For both top-quark modelling approaches one observes that quadratic contributions can exceed the linear terms induced by the same operators. This is particularly apparent for the $\mathcal{O}_{tG}$, where the squared cross section contribution is enhanced by about 50%. It is important to note that the EFT expansion holds at the level of the Lagrangian and thus the amplitude. The dimension-8 contribution that would induce a separate $\mathcal{O}(\Lambda^{-4})$ term must be genuinely smaller than the leading dimension-6 term [95] in order for the expansion to be consistent. On the other hand, the hierarchy between the $\mathcal{O}(\Lambda^{-2})$ and the squared dimension-6 term individually is not clear a priori. The interference of the SM amplitude with the SMEFT one can be destructive for various reasons, such as restrictions due to their helicity structure or their different distributions within the phase-space. For example, high p_T regions are enhanced in effective theories compared to the SM, while low p_T regimes are suppressed, because power-counting dictates a scaling behavior as $(E/\Lambda)^n$. The squared dimension-6 term is, of course, unrestricted by such interference effects and can thus appear bigger in practice. A consistent truncation of the prediction on the level of the observables would require the inclusion of the dimension-8 operators, which are not considered here. However, we still provide results for the quadratic dimension-6 terms as they often clarify the behavior of the SMEFT amplitudes, especially in differential distributions. Furthermore, experimental studies often quote results that also include quadratic terms, even when the dimension-8 term was not included. We remind the reader that the parametrization of σ_i and σ_{ij} is such that they are effectively computed using a NP scale of $\Lambda = 1\,\text{TeV}$. The validity of the EFT expansion naturally improves as the value of Λ increases or equivalently when

$pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ (NWA)					
	σ^{LO}[fb]	$\sigma^{\text{LO}}/\sigma^{\text{LO}}_{\text{SM}}$	σ^{NLO}[fb]	$\sigma^{\text{NLO}}/\sigma^{\text{NLO}}_{\text{SM}}$	$\mathcal{K}$
σ_{SM}	$2.2170^{+0.6593\,+0.0000}_{-0.4744\,-0.0000}$	$1.000^{+0.000\,+0.000}_{-0.000\,-0.000}$	$2.7425^{+0.0764\,+0.0000}_{-0.2062\,-0.0000}$	$1.000^{+0.000\,+0.000}_{-0.000\,-0.000}$	1.24
$\sigma_{t\phi}$	$-0.2713^{+0.0581\,+0.0121}_{-0.0807\,-0.0139}$	$-0.122^{+0.000\,+0.005}_{-0.000\,-0.006}$	$-0.3476^{+0.0286\,+0.0042}_{-0.0145\,-0.0055}$	$-0.127^{+0.001\,+0.002}_{-0.002\,-0.002}$	1.28
$\sigma_{\phi G}$	$1.5236^{+0.4669\,+0.1970}_{-0.3338\,-0.1630}$	$0.687^{+0.006\,+0.089}_{-0.005\,-0.074}$	$2.4696^{+0.3113\,+0.1233}_{-0.3132\,-0.1019}$	$0.900^{+0.086\,+0.045}_{-0.050\,-0.037}$	1.62
σ_{tG}	$2.2782^{+0.6917\,+0.0093}_{-0.4957\,-0.0119}$	$1.028^{+0.006\,+0.004}_{-0.006\,-0.005}$	$2.8064^{+0.0594\,+0.0143}_{-0.2059\,-0.0169}$	$1.023^{+0.002\,+0.005}_{-0.007\,-0.006}$	1.23
σ_{tW}	$0.6911^{+0.2059\,+0.0106}_{-0.1480\,-0.0114}$	$0.312^{+0.000\,+0.005}_{-0.000\,-0.005}$	$0.8619^{+0.0267\,+0.0030}_{-0.0663\,-0.0045}$	$0.314^{+0.001\,+0.001}_{-0.001\,-0.002}$	1.25
$\sigma_{t\phi,t\phi}$	$0.0083^{+0.0025\,+0.0009}_{-0.0018\,-0.0007}$	$0.004^{+0.000\,+0.000}_{-0.000\,-0.000}$	$0.0110^{+0.0006\,+0.0004}_{-0.0010\,-0.0003}$	$0.004^{+0.000\,+0.000}_{-0.000\,-0.000}$	1.33
$\sigma_{t\phi,\phi G}$	$-0.0932^{+0.0204\,+0.0137}_{-0.0286\,-0.0175}$	$-0.042^{+0.000\,+0.006}_{-0.000\,-0.008}$	$-0.1552^{+0.0203\,+0.0099}_{-0.0209\,-0.0118}$	$-0.057^{+0.003\,+0.004}_{-0.006\,-0.004}$	1.66
$\sigma_{t\phi,tG}$	$-0.1394^{+0.0303\,+0.0057}_{-0.0423\,-0.0064}$	$-0.063^{+0.000\,+0.003}_{-0.000\,-0.003}$	$-0.1778^{+0.0143\,+0.0013}_{-0.0063\,-0.0017}$	$-0.065^{+0.000\,+0.000}_{-0.001\,-0.001}$	1.28
$\sigma_{t\phi,tW}$	$-0.0423^{+0.0091\,+0.0013}_{-0.0126\,-0.0014}$	$-0.019^{+0.000\,+0.001}_{-0.000\,-0.001}$	$-0.0546^{+0.0046\,+0.0004}_{-0.0024\,-0.0006}$	$-0.020^{+0.000\,+0.000}_{-0.000\,-0.000}$	1.29
$\sigma_{\phi G,\phi G}$	$0.6815^{+0.2303\,+0.1524}_{-0.1607\,-0.1150}$	$0.307^{+0.012\,+0.069}_{-0.010\,-0.052}$	$1.4102^{+0.2697\,+0.1654}_{-0.2324\,-0.1349}$	$0.514^{+0.082\,+0.060}_{-0.050\,-0.049}$	2.07
$\sigma_{\phi G,tG}$	$1.3935^{+0.4573\,+0.1233}_{-0.3216\,-0.1066}$	$0.629^{+0.018\,+0.056}_{-0.016\,-0.048}$	$2.3829^{+0.3193\,+0.0754}_{-0.3227\,-0.0635}$	$0.869^{+0.090\,+0.027}_{-0.057\,-0.023}$	1.71
$\sigma_{\phi G,tW}$	$0.2381^{+0.0731\,+0.0263}_{-0.0522\,-0.0222}$	$0.107^{+0.001\,+0.012}_{-0.001\,-0.010}$	$0.3881^{+0.0497\,+0.0166}_{-0.0496\,-0.0136}$	$0.142^{+0.014\,+0.006}_{-0.008\,-0.005}$	1.63
$\sigma_{tG,tG}$	$2.8426^{+0.9654\,+0.0530}_{-0.6738\,-0.0588}$	$1.282^{+0.051\,+0.024}_{-0.045\,-0.027}$	$4.1576^{+0.2686\,+0.0387}_{-0.4342\,-0.0484}$	$1.516^{+0.054\,+0.014}_{-0.048\,-0.018}$	1.46
$\sigma_{tG,tW}$	$0.3559^{+0.1082\,+0.0073}_{-0.0775\,-0.0081}$	$0.161^{+0.001\,+0.003}_{-0.001\,-0.004}$	$0.4416^{+0.0106\,+0.0039}_{-0.0331\,-0.0051}$	$0.161^{+0.000\,+0.001}_{-0.001\,-0.002}$	1.24
$\sigma_{tW,tW}$	$0.0718^{+0.0214\,+0.0022}_{-0.0154\,-0.0024}$	$0.032^{+0.000\,+0.001}_{-0.000\,-0.001}$	$0.0895^{+0.0027\,+0.0006}_{-0.0068\,-0.0009}$	$0.033^{+0.000\,+0.000}_{-0.000\,-0.000}$	1.25

Table 6.4: *Integrated cross sections at LO and NLO for the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process with center-of-mass energy $\sqrt{s} = 13.6$ TeV. The SMEFT contributions are split into linear, quadratic and interference terms. The scales are set to $\mu_R = \mu_F = H_T/2$ as well as $\mu_{EFT} = m_t$ and the NNPDF3.1 PDF set is employed. The first uncertainty band corresponds to the 7-point scale variation of μ_R and μ_F, while the second one estimates the dependence on μ_{EFT}.*

the size of the Wilson coefficients is restricted below $\mathcal{O}(1)$. The results in this thesis can be trivially adjusted for higher Λ using Eq. (6.9) through a simple rescaling.

The $\mathcal{O}_{t\phi}$ and $\mathcal{O}_{\phi G}$ operators appear only in the production cross section of the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process. As one might expect we find that the integrated predictions relative to the SM for these operators are rather similar to the calculation with stable top-quarks. We conclude that the additional fiducial cuts applied to the NWA have only a small impact on these operators. This conclusion holds for both $\sigma_{t\phi}$ and $\sigma_{\phi G}$, as well their quadratic counterparts $\sigma_{t\phi,t\phi}$ and $\sigma_{\phi G,\phi G}$, respectively. A considerable change can be observed for the $\mathcal{O}_{tG}$ operator. The fiducial phase-space restriction leads to an increase of the σ_{tG} and $\sigma_{tG,tG}$ relative to the stable calculation at NLO. At LO the difference is less pronounced due to the absence of $\mathcal{O}_{tG}$ in the decay of the top. Additionally, a contribution of $\mathcal{O}_{tW}$ can of course only be studied within the calculation that includes the top-quark decays. Because this operators only enters inside the decays, we expect a smaller impact generally and consequently lower sensitivity.

The size of NLO QCD corrections, quantified by the $\mathcal{K}$-factors, can differ substantially between the SM and the SMEFT contributions. Starting with the fixed scale setup, in the NWA they lie in the range $1.25 - 1.99$ and for the stable calculation in $1.17 - 1.87$. The $\mathcal{K}$-factors for the SM are moderate at $\mathcal{K} = 1.27$ and $\mathcal{K} = 1.19$ for the stable and NWA calculations, respectively. For the dynamical scale choice

$pp \to t\bar{t}H + X$ (STABLE)					
	σ^{LO}[pb]	$\sigma^{\text{LO}}/\sigma^{\text{LO}}_{\text{SM}}$	σ^{NLO}[pb]	$\sigma^{\text{NLO}}/\sigma^{\text{NLO}}_{\text{SM}}$	$\mathcal{K}$
σ_{SM}	$0.4703^{+0.1440\,+0.0000}_{-0.1028\,-0.0000}$	$1.000^{+0.000\,+0.000}_{-0.000\,-0.000}$	$0.5673^{+0.0242\,+0.0000}_{-0.0479\,-0.0000}$	$1.000^{+0.000\,+0.000}_{-0.000\,-0.000}$	1.21
$\sigma_{t\phi}$	$-0.0576^{+0.0126\,+0.0026}_{-0.0176\,-0.0030}$	$-0.122^{+0.000\,+0.005}_{-0.000\,-0.006}$	$-0.0720^{+0.0066\,+0.0007}_{-0.0041\,-0.0010}$	$-0.127^{+0.001\,+0.001}_{-0.002\,-0.002}$	1.25
$\sigma_{\phi G}$	$0.3145^{+0.0992\,+0.0410}_{-0.0703\,-0.0339}$	$0.669^{+0.006\,+0.087}_{-0.005\,-0.072}$	$0.5067^{+0.0743\,+0.0230}_{-0.0699\,-0.0191}$	$0.893^{+0.089\,+0.041}_{-0.052\,-0.034}$	1.61
σ_{tG}	$0.4728^{+0.1483\,+0.0002}_{-0.1053\,-0.0006}$	$1.005^{+0.007\,+0.001}_{-0.006\,-0.001}$	$0.5691^{+0.0195\,+0.0026}_{-0.0467\,-0.0029}$	$1.003^{+0.003\,+0.005}_{-0.008\,-0.005}$	1.20
$\sigma_{t\phi,t\phi}$	$0.0018^{+0.0005\,+0.0002}_{-0.0004\,-0.0002}$	$0.004^{+0.000\,+0.000}_{-0.000\,-0.000}$	$0.0023^{+0.0002\,+0.0001}_{-0.0002\,-0.0001}$	$0.004^{+0.000\,+0.000}_{-0.000\,-0.000}$	1.30
$\sigma_{t\phi,\phi G}$	$-0.0192^{+0.0043\,+0.0028}_{-0.0061\,-0.0036}$	$-0.041^{+0.000\,+0.006}_{-0.000\,-0.008}$	$-0.0319^{+0.0045\,+0.0019}_{-0.0050\,-0.0022}$	$-0.056^{+0.004\,+0.003}_{-0.006\,-0.004}$	1.66
$\sigma_{t\phi,tG}$	$-0.0289^{+0.0064\,+0.0013}_{-0.0091\,-0.0014}$	$-0.062^{+0.000\,+0.003}_{-0.000\,-0.003}$	$-0.0362^{+0.0032\,+0.0002}_{-0.0018\,-0.0003}$	$-0.064^{+0.000\,+0.000}_{-0.001\,-0.001}$	1.25
$\sigma_{\phi G,\phi G}$	$0.1380^{+0.0481\,+0.0309}_{-0.0332\,-0.0233}$	$0.293^{+0.012\,+0.066}_{-0.010\,-0.050}$	$0.2914^{+0.0632\,+0.0333}_{-0.0518\,-0.0273}$	$0.514^{+0.086\,+0.059}_{-0.052\,-0.048}$	2.11
$\sigma_{\phi G,tG}$	$0.2868^{+0.0974\,+0.0262}_{-0.0678\,-0.0225}$	$0.610^{+0.019\,+0.056}_{-0.017\,-0.048}$	$0.4953^{+0.0777\,+0.0147}_{-0.0734\,-0.0124}$	$0.873^{+0.096\,+0.026}_{-0.061\,-0.022}$	1.73
$\sigma_{tG,tG}$	$0.5246^{+0.1811\,+0.0092}_{-0.1258\,-0.0103}$	$1.116^{+0.044\,+0.020}_{-0.039\,-0.022}$	$0.7525^{+0.0582\,+0.0070}_{-0.0839\,-0.0087}$	$1.326^{+0.044\,+0.012}_{-0.039\,-0.015}$	1.43

Table 6.5: *Integrated cross sections at LO and NLO for the $pp \to t\bar{t}H + X$ process with center-of-mass energy $\sqrt{s} = 13.6$ TeV. The SMEFT contributions are split into linear, quadratic and interference terms. The scales are set to $\mu_R = \mu_F = H_T/2$ as well as $\mu_{EFT} = m_t$ and the NNPDF3.1 PDF set is employed. The first uncertainty band corresponds to the 7-point scale variation of μ_R and μ_F, while the second one estimates the dependence on μ_{EFT}.*

we observe a slight increase in the $\mathcal{K}$-factors of about 5% in the NWA and about 10% in the stable calculation. The size of the NLO corrections itself is only LO accurate, which makes the difference between the scale setups unremarkable. The importance of NLO corrections is further stressed by the size of the uncertainties. Both the traditional scale uncertainties and the μ_{EFT} uncertainties are substantially reduced for both the SM and the SMEFT contributions. This conclusion holds in both computational setups. The only exception is the σ_{tG} term, which sees an increase in the latter type of uncertainty. This behavior is even more pronounced in the stable calculation compared to the NWA. However, this effect does not directly translate to the interference- and squared SMEFT contributions, where a decrease in overall uncertainty can always be observed. The degree to which the uncertainties are reduced depends on the operator in question. Indeed, in many cases one of the two errors sees a larger relative reduction compared to the other. The size of the uncertainties similarly varies depending on the contribution in question. The dominant source of theoretical uncertainty at both LO and NLO is the traditional scale variation, which quantify the lack of additional terms in the expansion in α_s. Taking the ratio to the SM results reduces the scale variation errors, while leaving the μ_{EFT} related uncertainty unaffected. As a result, the two uncertainties become similar in magnitude. Indeed, at LO, the μ_{EFT} uncertainty can be even become dominant in some cases, such as $\sigma_{\phi G}$ and $\sigma_{\phi G,\phi G}$. The size of NLO uncertainties is generally reduced in the NWA compared to the stable calculation, which constitutes the first advantage of the consistent modelling of production and decay in this process. Furthermore, the access to a realistic phase-space selection and observables defined based on the fully decayed final state allows for better comparison with experimental results from ATLAS and CMS. The unfolding process to the full phase-space considered in calculations with stable top-quarks might be affected by

the inclusion of SMEFT operators and is thus of special interest.

Particularly noticeable is the large $\mathcal{K}$-factor for the $\mathcal{O}_{\phi G}$ linear and interference contributions, which are around 60-65% for both the NWA and stable setups. Even further, the NLO QCD corrections in the squared term $\sigma_{\phi G,\phi G}$ reach 90% (105%) and 100% (110%) for the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ and $pp \to t\bar{t}H + X$ calculations with the fixed (dynamical) scale settings, respectively. Additionally, the size of scale uncertainties at NLO is larger than for the other partial contributions in our calculations. These observations can be explained by new topologies that open up in the real-emission contributions at NLO. Representative diagrams are shown in Figure 6.4. One can observe that these topologies are not related to any Born diagram through QCD splitting. Consequently, these induce LO-precise contributions to our calculation starting at NLO, which cause both the relatively larger uncertainties and the large $\mathcal{K}$-factors in the $\mathcal{O}_{\phi G}$ induced terms.

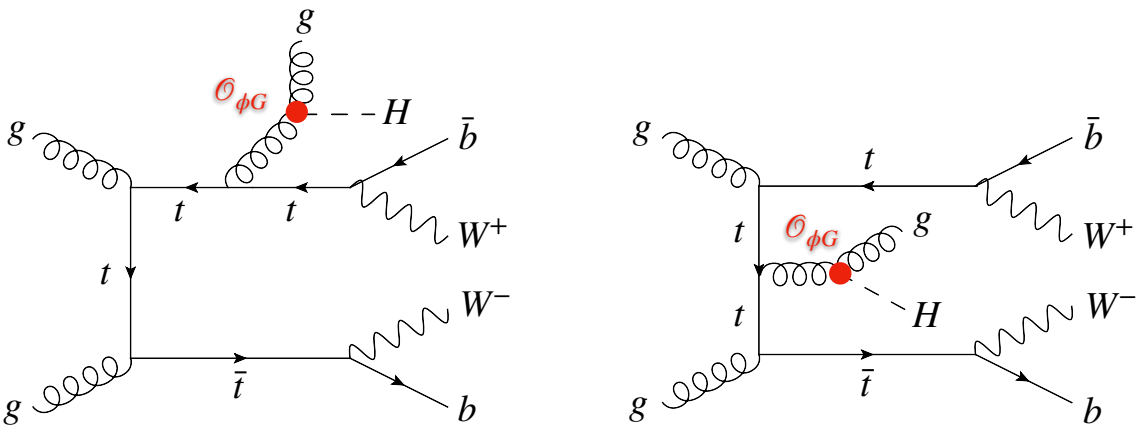


Figure 6.4: *Representative diagrams of new topologies opening up in the real-emission contribution at NLO QCD within in the $C_{\phi G}$ partial cross section of the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process. The decays of the W bosons are suppressed for clarity. The same contributions can also be found in the $pp \to t\bar{t}H + X$ process. Taken from Ref. [3].*

We can substantiate this claim by imposing an additional restriction on the phase-space of the extra parton. To that end, we require a jet-veto, such that $p_T(j) < p_T^{veto}$, which leaves the singular regions of the real-emission contribution unchanged while reducing additional hard radiation. Tables 6.6 and 6.7 compare the corresponding NLO cross sections in the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process with no jet-veto and jet-veto values of $p_T^{veto} = 100\,\text{GeV}$ and $p_T^{veto} = 150\,\text{GeV}$ for the fixed and dynamical scale setups, respectively. We focus only on the $\sigma_{\phi G}$ and $\sigma_{\phi G,\phi G}$ results for clarity, but also provide the SM cross section for comparison. The dominant source of theoretical errors remains the μ_R/μ_F scale uncertainty even when applying the jet-veto. However, for $\sigma_{\phi G}$ using the fixed scale setup, the size is reduced from 12% to 8% and 5% when applying $p_T^{veto} = 150\,\text{GeV}$ and $p_T^{veto} = 100\,\text{GeV}$, respectively. The $\mathcal{K}$-factors are reduced from 56% to 28% and 12% as well. Similarly, in the dynamical scale setup a reduction of the scale uncertainties from 12% to 9% and 7% can be observed. The case of $\sigma_{\phi G,\phi G}$ is even clearer. The size of the scale variation for fixed (dynamical) setup is reduced from 16% (19%) to 11% (13%) and 8% (11%) for jet-vetos of 150 GeV and 100 GeV, while the size of the NLO QCD corrections is reduced from 87% (107%) to 49% (75%) and 30% (58%). On the other hand, the uncertainty

of the SM prediction with fixed scales increases from 7% to 10% to 20% as the phase-space becomes more restricted. This highlights that the perturbative stability of the calculation is generally reduced when imposing a jet-veto. Similar behavior can be found for the $\sigma_{\phi G}$ contributions, but here the reduction of LO-like real-emission topologies competes with the decreased perturbative stability of the SM amplitude. We note, however, that we only study the inclusion of a jet-veto to highlight that such real-emission contributions play a large role in the $\mathcal{O}_{\phi G}$ terms. The dynamical scale setup is generally more stable when imposing a jet-veto. The SM uncertainty increases only to 13% instead of 20% when imposing $p_T(j) < 100\,\text{GeV}$, while the $p_T^{veto} = 150\,\text{GeV}$ leaves the size of the theoretical uncertainty intact at about 7%. This conclusion is even more relevant on the differential level, where predictions with fixed scale choices can quickly become unstable in specific phase-space regions when employing a jet-veto.

p_T^{veto} [GeV]		σ^{LO} [fb]	σ^{NLO} [fb]	$\mathcal{K}$
-	σ_{SM}	$2.3166^{+0.7068\,+0.0000}_{-0.5044\,-0.0000}$	$2.7680^{+0.0560\,+0.0000}_{-0.1898\,-0.0000}$	1.19
	$\sigma_{\phi G}$	$1.6262^{+0.5137\,+0.2089}_{-0.3638\,-0.1730}$	$2.5292^{+0.2865\,+0.1154}_{-0.3098\,-0.0954}$	1.56
	$\sigma_{\phi G,\,\phi G}$	$0.8178^{+0.2987\,+0.1799}_{-0.2032\,-0.1361}$	$1.5299^{+0.2443\,+0.1528}_{-0.2377\,-0.1277}$	1.87
150	σ_{SM}	$2.3166^{+0.7068\,+0.0000}_{-0.5044\,-0.0000}$	$2.2694^{+0.0171\,+0.0000}_{-0.2326\,-0.0000}$	0.98
	$\sigma_{\phi G}$	$1.6262^{+0.5137\,+0.2089}_{-0.3638\,-0.1730}$	$2.0748^{+0.0601\,+0.0583}_{-0.1642\,-0.0486}$	1.28
	$\sigma_{\phi G,\,\phi G}$	$0.8178^{+0.2987\,+0.1799}_{-0.2032\,-0.1361}$	$1.2207^{+0.0738\,+0.0840}_{-0.1339\,-0.0761}$	1.49
100	σ_{SM}	$2.3166^{+0.7068\,+0.0000}_{-0.5044\,-0.0000}$	$1.9589^{+0.0652\,+0.0000}_{-0.3863\,-0.0000}$	0.85
	$\sigma_{\phi G}$	$1.6262^{+0.5137\,+0.2089}_{-0.3638\,-0.1730}$	$1.8239^{+0.0032\,+0.0259}_{-0.0863\,-0.0220}$	1.12
	$\sigma_{\phi G,\,\phi G}$	$0.8178^{+0.2987\,+0.1799}_{-0.2032\,-0.1361}$	$1.0616^{+0.0209\,+0.0483}_{-0.0815\,-0.0494}$	1.30

Table 6.6: *Cross section results for the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process at LO and NLO QCD computed in the NWA with center-of-mass energy $\sqrt{s} = 13.6$ TeV. Results are provided with two values of the jet-veto, given by $p_T^{veto} = 150$ GeV and $p_T^{veto} = 100$ GeV. For comparison the unrestricted result is also shown. The terms σ_{SM}, $\sigma_{\phi G}$ and $\sigma_{\phi G,\,\phi G}$ are shown for each jet-veto setting. The last column also presents the $\mathcal{K}$-factors. The scales are set to $\mu_R = \mu_F = m_t + m_H/2$ as well as $\mu_{EFT} = m_t$ and the NNPDF3.1 PDF set is used. The first uncertainty band refers to the variation of μ_R and μ_F and the second one to μ_{EFT}. Taken from Ref. [3].*

To further study the differences of the operator contributions in the $pp \to t\bar{t}H+X$ and $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ processes, we define the double ratio

$$\mathcal{D}_i \equiv \frac{\sigma_{\text{i}}/\sigma_{\text{SM}}}{\sigma_{\text{i}}^{\text{NWA}}/\sigma_{\text{SM}}^{\text{NWA}}}\,, \tag{6.20}$$

where the denominator was chosen so that terms of different orders of magnitude

p_T^{veto} [GeV]		σ^{LO} [fb]	σ^{NLO} [fb]	$\mathcal{K}$
-	σ_{SM}	$2.2170^{+0.6593\,+0.0000}_{-0.4744\,-0.0000}$	$2.7425^{+0.0764\,+0.0000}_{-0.2062\,-0.0000}$	1.24
	$\sigma_{\phi G}$	$1.5236^{+0.4669\,+0.1970}_{-0.3338\,-0.1630}$	$2.4696^{+0.3113\,+0.1233}_{-0.3132\,-0.1019}$	1.62
	$\sigma_{\phi G,\,\phi G}$	$0.6815^{+0.2303\,+0.1524}_{-0.1607\,-0.1150}$	$1.4102^{+0.2697\,+0.1654}_{-0.2324\,-0.1349}$	2.07
150	σ_{SM}	$2.2170^{+0.6593\,+0.0000}_{-0.4744\,-0.0000}$	$2.3184^{+0.0094\,+0.0000}_{-0.1326\,-0.0000}$	1.05
	$\sigma_{\phi G}$	$1.5236^{+0.4669\,+0.1970}_{-0.3338\,-0.1630}$	$2.0952^{+0.1261\,+0.0760}_{-0.1978\,-0.0631}$	1.38
	$\sigma_{\phi G,\phi G}$	$0.6815^{+0.2303\,+0.1524}_{-0.1607\,-0.1150}$	$1.1938^{+0.1599\,+0.1163}_{-0.1644\,-0.0982}$	1.75
100	σ_{SM}	$2.2170^{+0.6593\,+0.0000}_{-0.4744\,-0.0000}$	$2.0348^{+0.0264\,+0.0000}_{-0.2686\,-0.0000}$	0.92
	$\sigma_{\phi G}$	$1.5236^{+0.4669\,+0.1970}_{-0.3338\,-0.1630}$	$1.8717^{+0.0372\,+0.0469}_{-0.1299\,-0.0392}$	1.23
	$\sigma_{\phi G,\phi G}$	$0.6815^{+0.2303\,+0.1524}_{-0.1607\,-0.1150}$	$1.0744^{+0.1002\,+0.0890}_{-0.1273\,-0.0778}$	1.58

Table 6.7: *Cross section results for the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H+X$ process at LO and NLO QCD computed in the NWA with center-of-mass energy $\sqrt{s} = 13.6$ TeV. Results are provided with two values of the jet-veto, given by $p_T^{veto} = 150$ GeV and $p_T^{veto} = 100$ GeV. For comparison the unrestricted result is also shown. The terms σ_{SM}, $\sigma_{\phi G}$ and $\sigma_{\phi G,\,\phi G}$ are shown for each jet-veto setting. The last column also presents the $\mathcal{K}$-factors. The scales are set to $\mu_R = \mu_F = H_T/2$ as well as $\mu_{EFT} = m_t$ and the NNPDF3.1 PDF set is used. The first uncertainty band refers to the variation of μ_R and μ_F and the second one to μ_{EFT}.*

are always centered at 1. We compute the uncertainties of $\mathcal{D}_i$ purely by rescaling the uncertainty of the ratio in the numerator. With this definition, $\mathcal{D}_i$ provides a measure for the relative difference of the NWA to the stable top quark calculations and the error bars highlight if the predictions can be expected to agree within their uncertainties. Figure 6.5 shows these double-ratios at LO (upper plot) and NLO (lower plot) for $\mu_R = \mu_F = m_t + m_H/2$, where the blue values correspond to the stable calculation and the red ones to the NWA. The inner error bars correspond to traditional scale uncertainties, while the outer ones are related to the value of the μ_{EFT} definition. We also present the $\mathcal{O}_{tW}$ results to allow for a visual comparison of the uncertainties with other operators even though that the stable calculation does not include this contribution. Starting with LO we can see that differences mostly arise in the σ_{tG} and $\sigma_{tG,tG}$. We remind the reader that at LO the $\mathcal{O}_{tG}$ operator only enters in the production part of the process. Therefore, we can conclude that this difference is the result of the fiducial phase-space cuts imposed in the NWA. To supplement the discussion we also provide the same double-ratios at LO with no phase-space cuts applied to the NWA calculation in Figure 6.6. As expected no difference can be observed for $\mathcal{O}_{tG}$ in this case.

Returning to the NLO double-ratio plot shown in Figure 6.5, we can make similar observations. While the smaller uncertainties generally increase the significance of

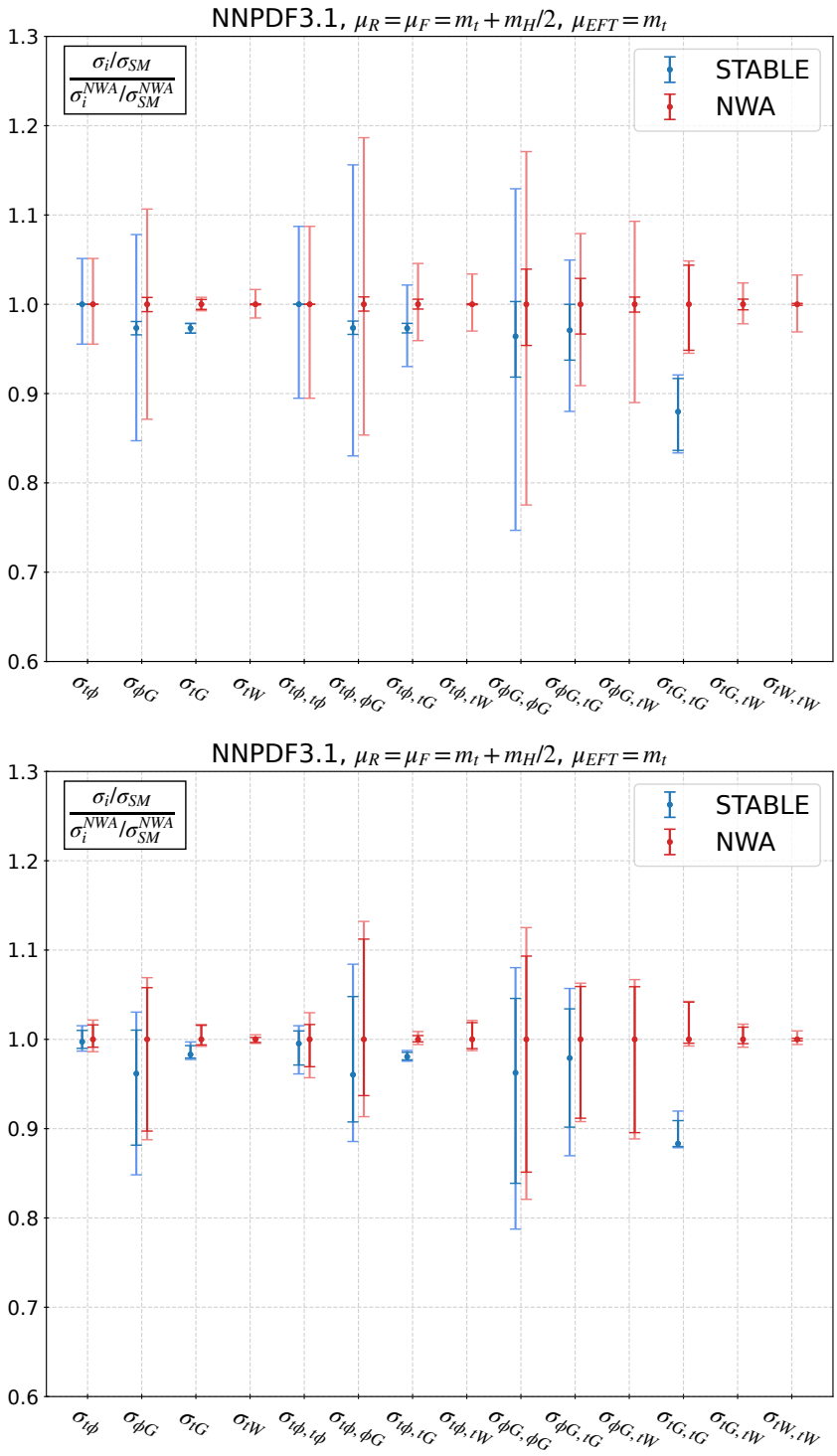


Figure 6.5: *Double ratio of integrated cross section predictions defined as* $\mathcal{D}_i \equiv \frac{\sigma_i/\sigma_{SM}}{\sigma_i^{NWA}/\sigma_{SM}^{NWA}}$ *at LO (upper plot) and NLO (lower plot). The calculations with stable top-quarks and top-quark decays in the NWA are shown in blue and red, respectively. Linear, quadratic and interference terms are shown. All results are evaluated using the NNPDF3.1 PDF set and* $\mu_R = \mu_F = m_t + m_H/2$ *as well as* $\mu_{EFT} = m_t$ *at* $\sqrt{s} = 13.6$ *TeV. The inner error bands correspond to the 7-point scale variations of* μ_R *and* μ_F*, while the outer also includes the uncertainty related to the definition of* μ_{EFT}*. All uncertainty bands are computed based on the numerator in* $\mathcal{D}_i$ *and only rescaled by the denominator. Taken from Ref. [3].*

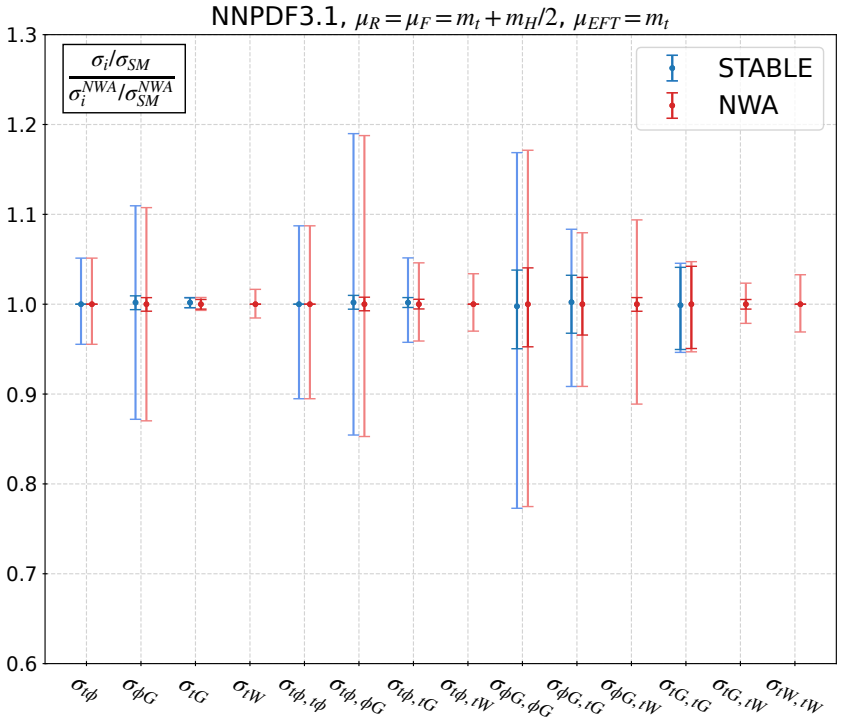


Figure 6.6: *Same as the LO (upper) panel of Figure 6.5, but with a fully inclusive phase-space selection. Taken from Ref. [3].*

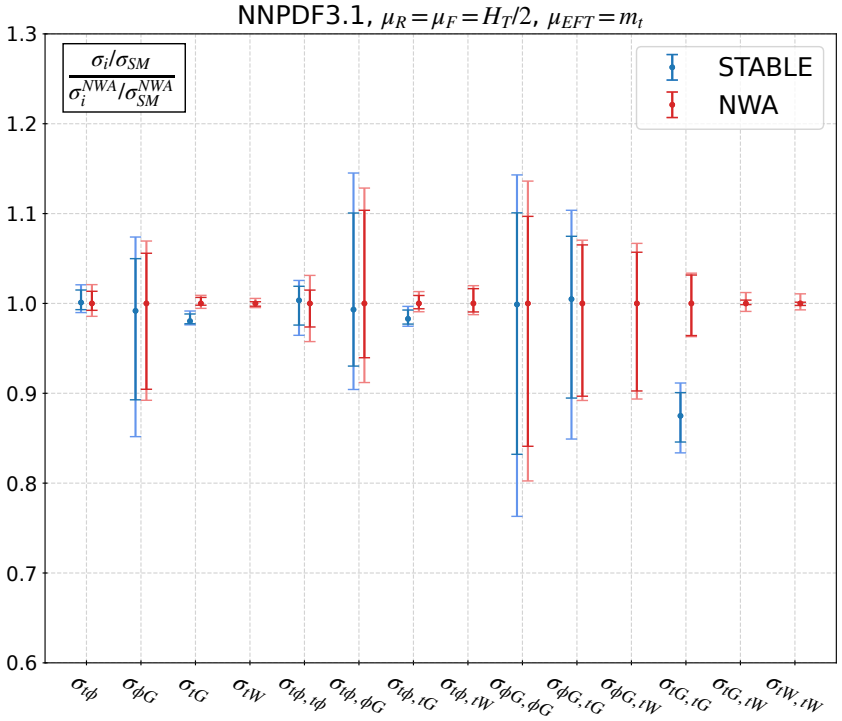


Figure 6.7: *Same as the NLO panel of Figure 6.5, but with the dynamical scale choice $\mu_R = \mu_F = H_T/2$ and $\mu_{EFT} = m_t$.*

the differences in $\mathcal{O}_{tG}$ induced contributions, we mostly see good agreement for $\mathcal{O}_{t\phi}$ and $\mathcal{O}_{\phi G}$ on the integrated level. However, the $\mathcal{O}_{tG}$ contributions are significantly changed, and effects from fiducial phase-space cuts are needed to correctly describe them. We note that the fiducial cuts imposed in this study are very inclusive and an analyses in specific phase-space regions might see much larger effects. This is particularly relevant to new physics searches, which often analyse high energy regimes, where the SM contributions are relatively suppressed.

To briefly touch on the dynamical scale choice, we present the double-ratio plot again at NLO only with $\mu_R = \mu_F = H_T/2$ and $\mu_{EFT} = m_t$ in Figure 6.7. Due to the slight increase in the theoretical uncertainty for the $\mathcal{O}_{tG}$ operator, we observe slightly better agreement in the corresponding contributions. However, the conclusions remain the same. Contributions proportional to C_{tG} are sensitive to the phase-space selection with $\sigma_{tG,tG}$ differing by more than 2 standard deviations. Similarly, the linear contribution is slightly outside the uncertainty estimate. On the other hand $\mathcal{O}_{t\phi}$ and $\mathcal{O}_{\phi G}$ are well described in the stable top-quark approximation.

We conclude that the calculation of QCD corrections is generally important even at the integrated level to control theoretical uncertainties, while the calculation of top-quark decays and proper phase-space selections can strongly influence the extracted results for specific operators even if they don't appear in the decay process directly. Lastly, we note that some operators are simply not accessible in calculations with stable top-quarks and need to be studied within a more complete description of the processes.

6.3 Differential distributions

The impact of NLO corrections can be substantially increased in specific phase-space regions. Furthermore, new-physics searches are often particularly interested in regions such as high-p_T tails, where NP effects are expected to occur. In this section, we investigate the impact of NLO QCD corrections on various kinematic distributions of the Higgs boson, leptons and b-jets and compare the shape of the various SMEFT contributions. In particular, we take note when the shapes of the SM and the SMEFT contributions differ significantly, as these regions are most likely to be sensitive to experimental tests. We focus on the calculation in the NWA, that is the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process, and provide a multitude of dimensionless and dimensionful observables. We have seen on the level of the integrated cross sections that the dominant uncertainty originates from the variation of the renormalization and factorization scales. This conclusion generalizes to the differential level for all observables that we study in this thesis. We briefly demonstrate this fact by presenting the size of the theoretical uncertainties on the transverse momentum distribution of the Higgs boson in Figure 6.8. Each panel corresponds to a different contribution at NLO QCD parametrized according to Eq. (6.9). The error due to μ_R and μ_F exceeds the μ_{EFT} variation by a factor of 3 or more for $\sigma_{t\phi}$, σ_{tG}, and σ_{tW}. For $\sigma_{\phi G}$, we observe at least a factor of 2 difference in the size of the two uncertainties. The conclusion is the same as for the linear SMEFT contributions in the cases of $\sigma_{tG,tG}$ and $\sigma_{tW,tW}$. In contrast, the μ_{EFT} variation becomes more important for $\sigma_{t\phi,t\phi}$ and $\sigma_{\phi G,\phi G}$. However, we still find that the dominant contribution is given by the conventional scale variation. In the worst case, the theoretical uncertainty estimate increases from 20% to 25% after adding both terms in quadrature. To simplify

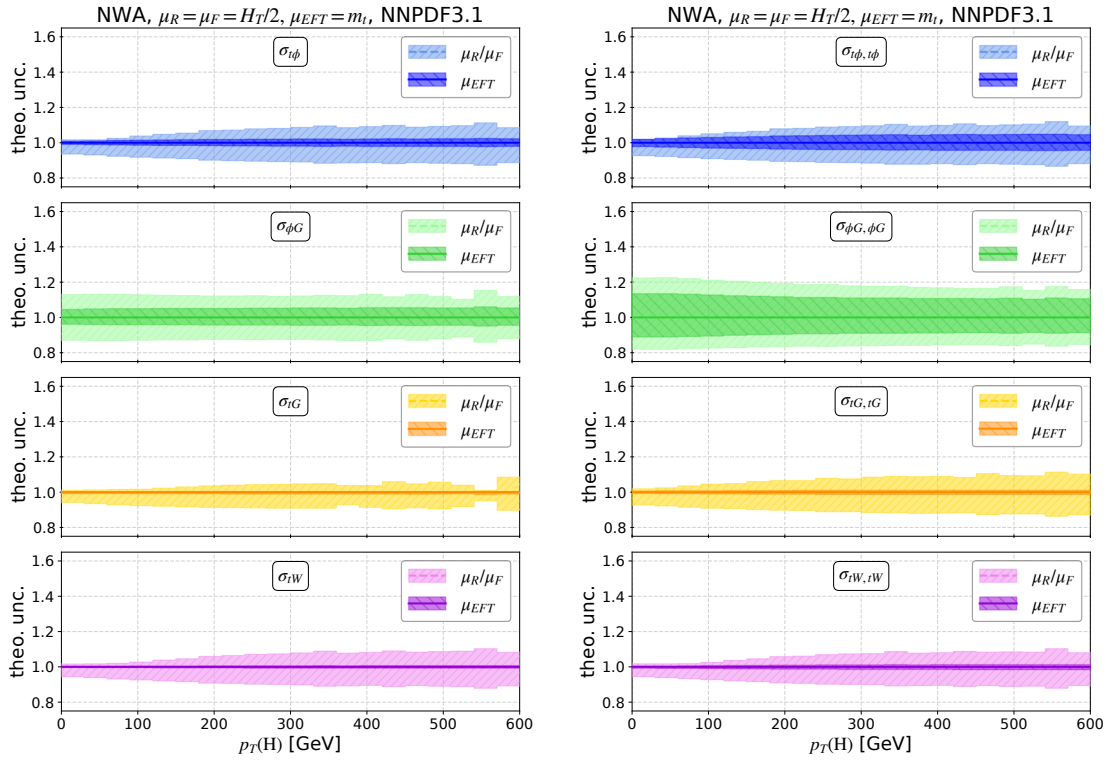


Figure 6.8: *Theoretical uncertainties of the $p_T(H)$ distribution due to the variation of μ_R/μ_F and μ_{EFT} at NLO QCD for the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process. The left column displays linear dimension-6 SMEFT contributions, while the right one contains the squared ones. All results are evaluated using the NNPDF3.1 PDF set and $\mu_R = \mu_F = H_T/2$ as well as $\mu_{EFT} = m_t$.*

the following discussion, we will thus always consider only the μ_R/μ_F variation to estimate the systematic uncertainty of the calculation. However, we do note that fits including squared dimension-6 terms must carefully check if this assumption is justified for all included operators.

We now return to the kinematical features of the SMEFT contributions. We present a series of differential distributions to highlight key observations. For each plot we provide the normalized distribution in the first panel. The left column corresponds to linear SMEFT terms σ_i and the right column to quadratic contributions σ_{ii}. The SM prediction is included in both columns to facilitate comparisons. The lower panels display differential $\mathcal{K}$-factors for each respective operator separately. In addition, scale uncertainties from the μ_R/μ_F variation are also shown. The distributions are evaluated with the NNPDF3.1 PDF set and the dynamical central scale choice $\mu_R = \mu_F = H_T/2$ with $\mu_{EFT} = m_t$. The dynamical scale choice typically stabilizes the perturbative expansion in harder phase-space regions, which we will comment on in the following sections as well.

We start by discussing the transverse momentum distribution of the Higgs boson, $p_T(H)$, presented in Figure 6.9. The kinematics of the Higgs boson can be used to disentangle the $\mathcal{O}_{\phi G}$ and $\mathcal{O}_{tG}$ contributions from the SM. This is true in the linear interference contribution, but significantly amplified in the squared term. In both cases these operators generate harder p_T spectra than the SM prediction. This effect can be understood by considering the Feynman rules for these operators.

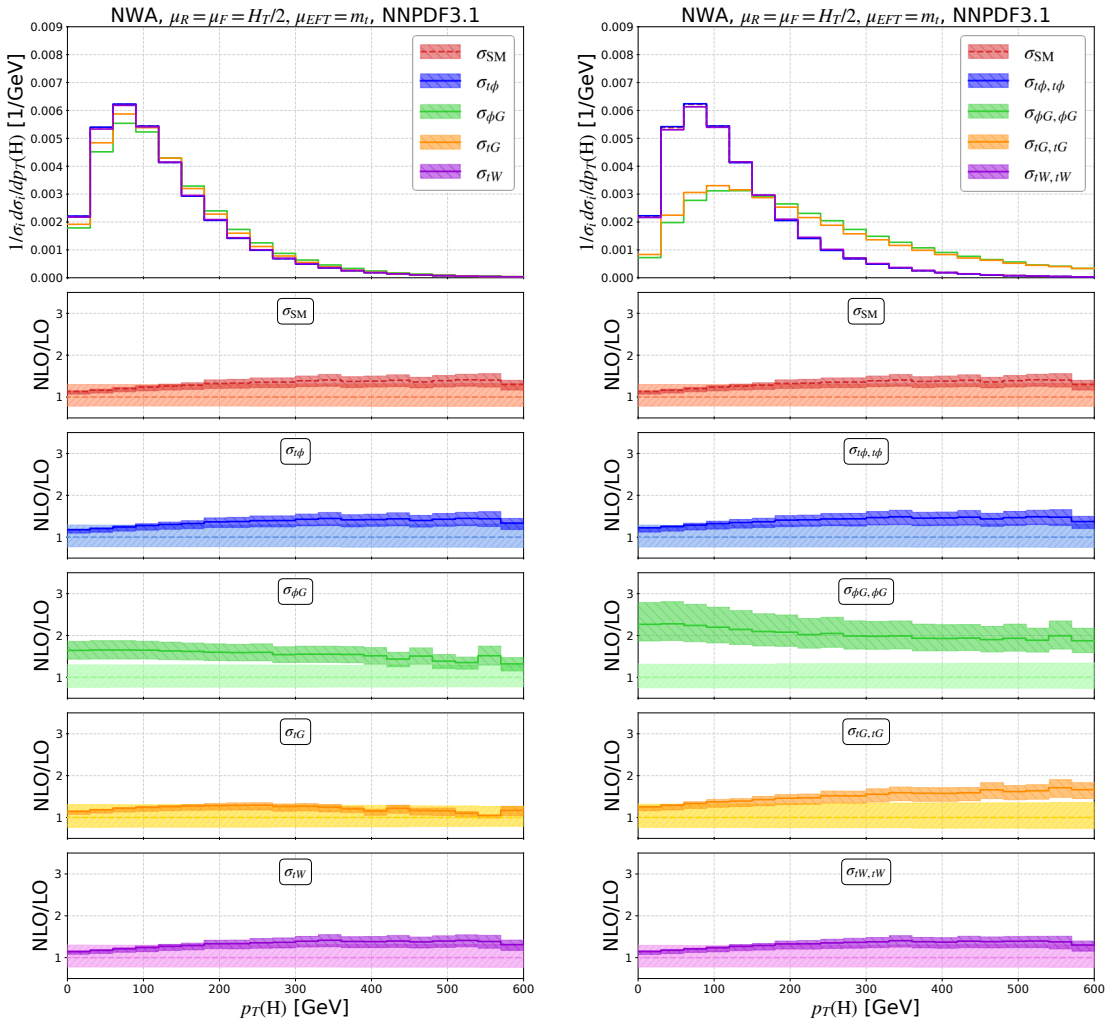


Figure 6.9: *Normalized differential cross section distributions in $p_T(H)$ for the $pp \to e^+\nu_e \mu^-\bar{\nu}_\mu b\bar{b} H + X$ process at NLO QCD. The first panel shows the SM predictions as well as the linear (squared) SMEFT contributions in the left (right) column. Differential $\mathcal{K}$-factors are presented in the lower panels for each contribution separately alongside their respective scale uncertainties. Uncertainties related to the choice of μ_{EFT} are not shown. All results are evaluated using the NNPDF3.1 PDF set and $\mu_R = \mu_F = H_T/2$ as well as $\mu_{EFT} = m_t$.*

Both of these generate couplings between the gluon and Higgs bosons that increase in strength with the momentum of the gluon, which in turn makes a stronger recoil of the Higgs more likely. The remaining operators produce distributions that look very similar to the SM and can not be distinguished using this observable. As already mentioned, the size of NLO QCD corrections varies for each operator. The $\sigma_{\phi G}$ and $\sigma_{\phi G,\phi G}$ are strongly enhanced at NLO in the low $p_T(H)$ region, due to the aforementioned new topologies that open up in the real-emission channels. They reach up to 70% and 130% for the linear and squared term, respectively.

One advantage of the calculation in the NWA is that we can study the kinematics of the top-quark decay products directly. The transverse momentum of the hardest b-jet $p_T(b_1)$ is presented in Figure 6.10. The picture is similar to the $p_T(H)$ case, albeit less pronounced. Differences arise mostly for $\mathcal{O}_{\phi G}$ and $\mathcal{O}_{tG}$, but the linear

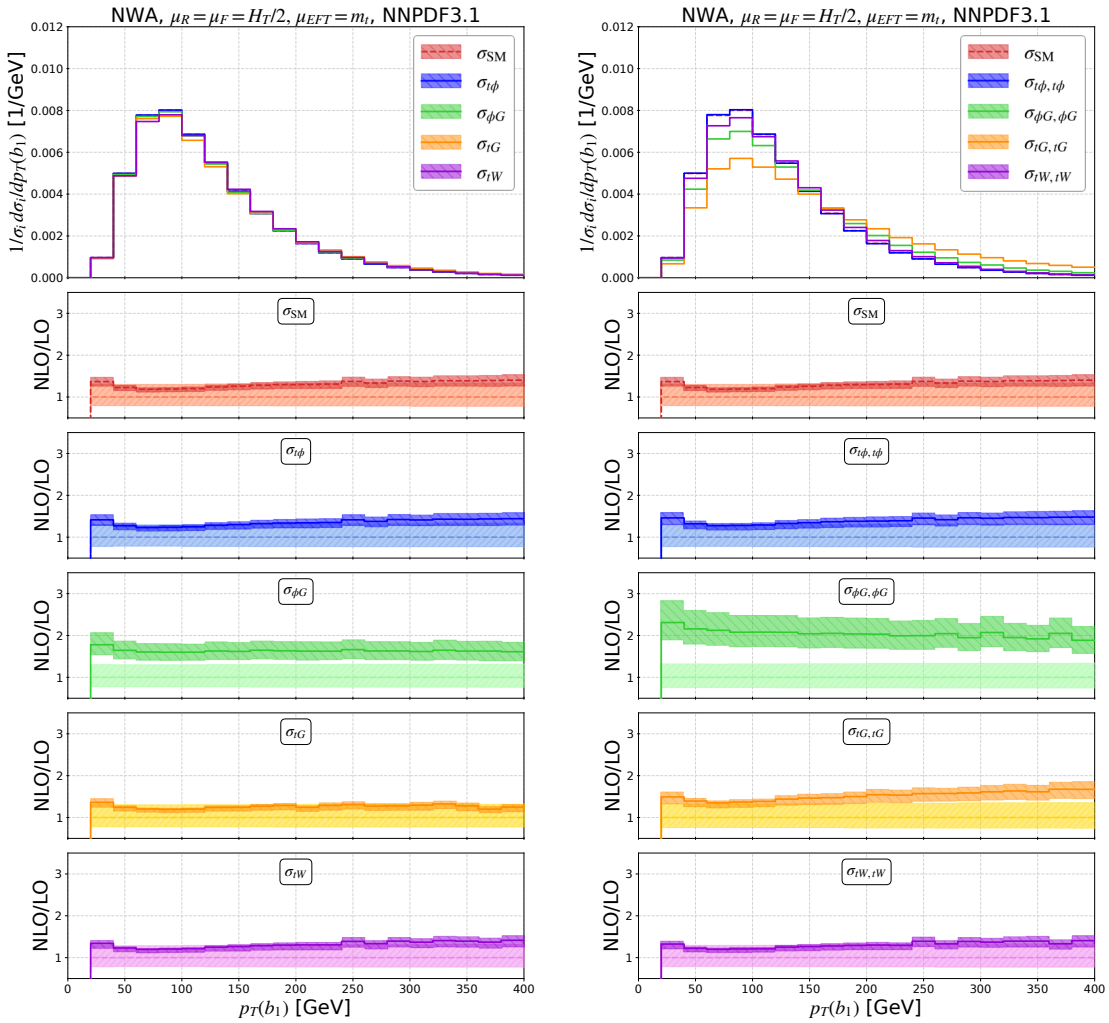


Figure 6.10: *Normalized differential cross section distributions in $p_T(b_1)$ for the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process at NLO QCD. The first panel shows the SM predictions as well as the linear (squared) SMEFT contributions in the left (right) column. Differential $\mathcal{K}$-factors are presented in the lower panels for each contribution separately alongside their respective scale uncertainties. Uncertainties related to the choice of μ_{EFT} are not shown. All results are evaluated using the NNPDF3.1 PDF set and $\mu_R = \mu_F = H_T/2$ as well as $\mu_{EFT} = m_t$.*

terms only show minimal differences. At the level of the squared terms one can still see the differences clearly. $\mathcal{K}$-factors are large for $\sigma_{\phi G}$ and $\sigma_{\phi G,\phi G}$ especially at the beginning of the spectrum, where they reach around 65% and 110%, respectively.

The kinematics of the leptonic decay products of the W boson are presented using the example of the transverse momentum distribution of the hardest lepton, $p_T(\ell_1)$, and the missing transverse momentum p_T^{miss}. We display these two distributions in Figures 6.11 and 6.12, respectively. The pattern for $\mathcal{O}_{\phi G}$ and $\mathcal{O}_{tG}$ in the $p_T(H)$ and $p_T(b_1)$ distributions generally also applies here. These operators favor harder spectra compared to the SM. Interestingly, these observables are also sensitive to effects from the $\mathcal{O}_{tW}$ operator. Shape differences relative to the SM typically range between 10% and 20% in large sections of the spectrum. As one would expect, these effects are stronger in the squared terms, but are still clearly visible in the linear

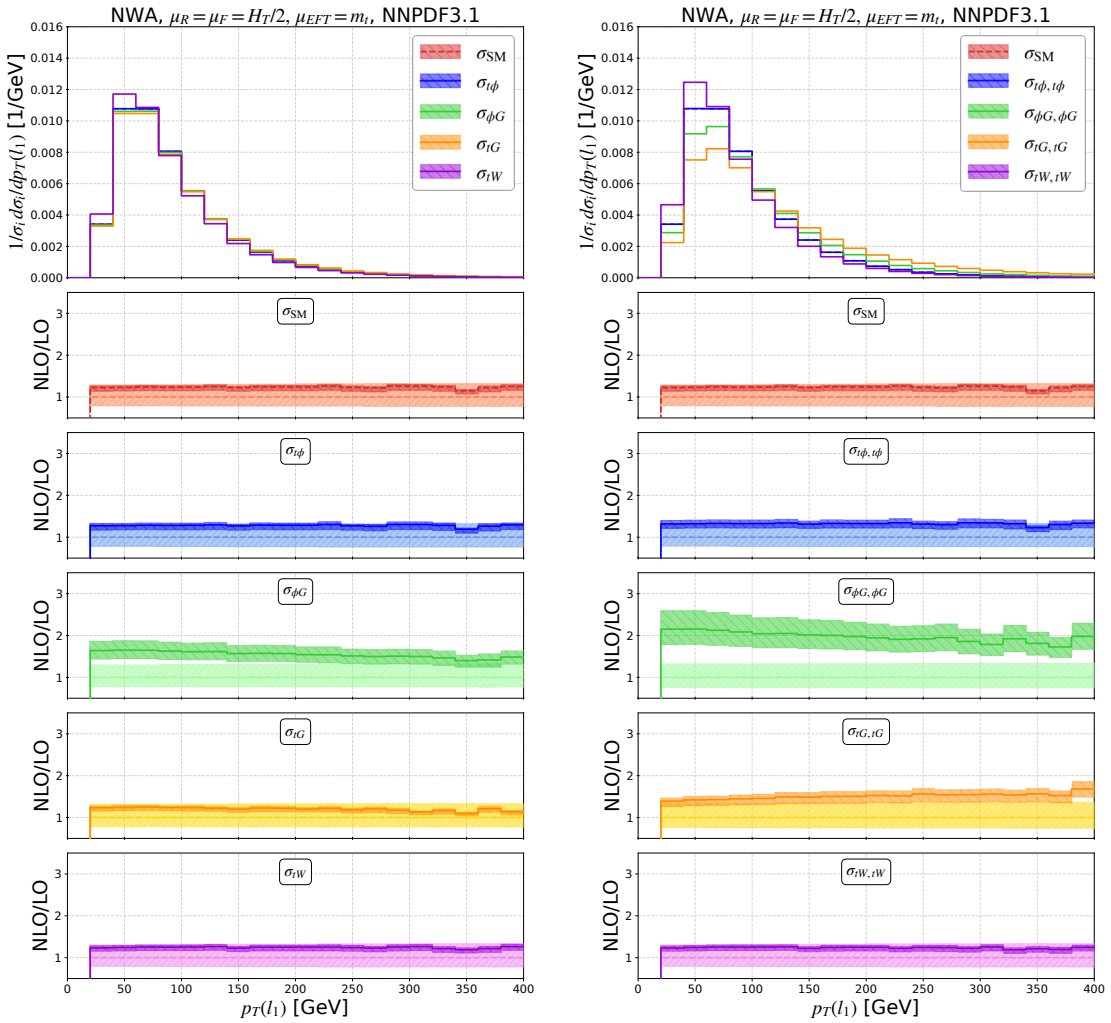


Figure 6.11: *Normalized differential cross section distributions in $p_T(\ell_1)$ for the $pp \to e^+\nu_e \mu^-\bar{\nu}_\mu b\bar{b} H + X$ process at NLO QCD. The first panel shows the SM predictions as well as the linear (squared) SMEFT contributions in the left (right) column. Differential $\mathcal{K}$-factors are presented in the lower panels for each contribution separately alongside their respective scale uncertainties. Uncertainties related to the choice of μ_{EFT} are not shown. All results are evaluated using the NNPDF3.1 PDF set and $\mu_R = \mu_F = H_T/2$ as well as $\mu_{EFT} = m_t$.*

one as well. We find that the $\mathcal{O}_{tW}$ contributions prefer softer transverse momenta in the W decay compared to the SM equivalent.

The scalar sum of the final state transverse momenta H_T is shown in Figure 6.13. The shape of the distribution shows similar features to the individual p_T spectra. Both $\mathcal{O}_{\phi G}$ and $\mathcal{O}_{tG}$ prefer harder momenta, which translates to a 15% reduction at the peak of the histogram and a 20% increase in the tail for these two operator contributions. At the level of squared terms this is further amplified. NLO QCD corrections remain mostly flat in the tail of the distribution due to the dynamical scale choice. The beginning of the histogram can show larger differences, because the inclusion of an additional jet in the real-emission contribution smears out the hard phase-space cut-off present at LO. With exception of $\mathcal{O}_{\phi G}$ the different contributions typically see corrections at about the size of the LO scale bands. In the case of $\sigma_{tG,tG}$

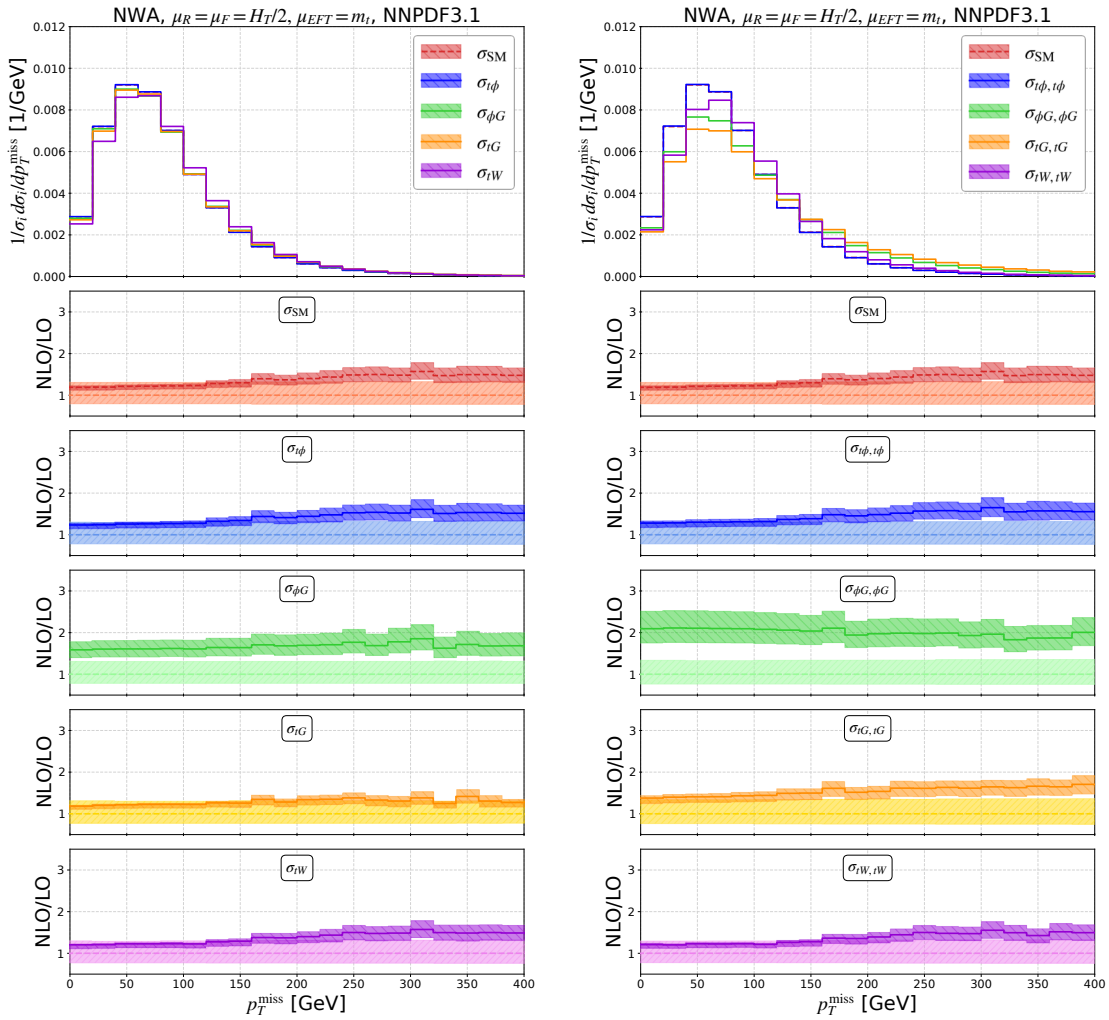


Figure 6.12: *Normalized differential cross section distributions in $p_T(miss)$ for the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process at NLO QCD. The first panel shows the SM predictions as well as the linear (squared) SMEFT contributions in the left (right) column. Differential $\mathcal{K}$-factors are presented in the lower panels for each contribution separately alongside their respective scale uncertainties. Uncertainties related to the choice of μ_{EFT} are not shown. All results are evaluated using the NNPDF3.1 PDF set and $\mu_R = \mu_F = H_T/2$ as well as $\mu_{EFT} = m_t$.*

one can find an additional enhancement, which leads to the correction being just outside the LO one. In the tail they reach about 50%, while the LO bands cover up to 30%. As expected $\mathcal{O}_{\phi G}$ sees large corrections at about 100%, which is inline with the integrated cross section.

Of particular interest for experimental analyses are angular observables, which can be sensitive to spin-correlations and other effects that might be modified in new physics contributions. We first discuss the distance of the two leptons in the rapidity-azimuthal angle plane $\Delta R(\ell_1, \ell_2)$ as well as the distribution of the azimuthal-angle between the leptons $\Delta\phi_{\ell_1,\ell_2}/\pi$, which are presented in Figure 6.14 and 6.15, respectively. The latter one is sensitive to spin-correlation effects, which makes it particularly interesting in the context of top-quark decays, which transfer their spin-structure to the decay products. The biggest SMEFT effects can be observed for

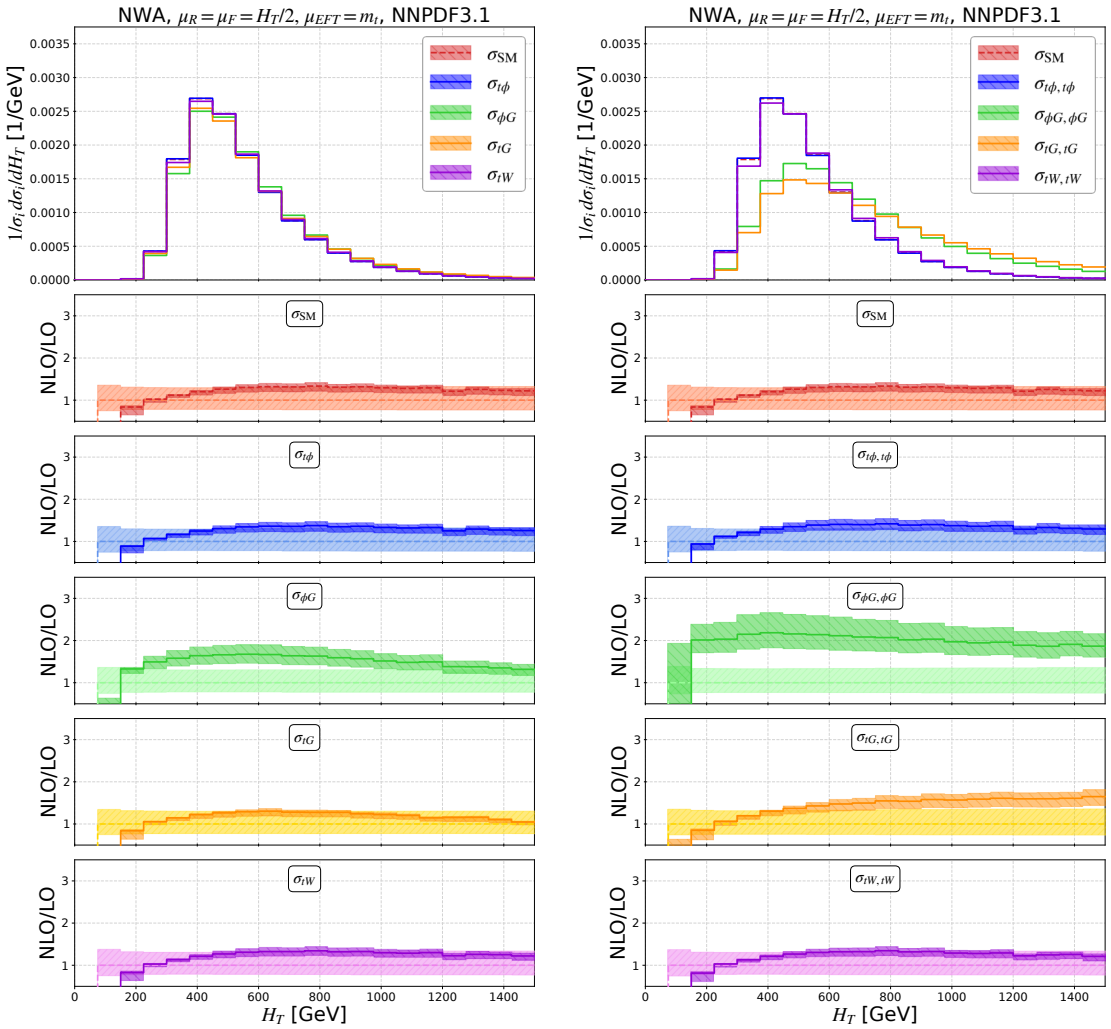


Figure 6.13: *Normalized differential cross section distributions in H_T for the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process at NLO QCD. The first panel shows the SM predictions as well as the linear (squared) SMEFT contributions in the left (right) column. Differential $\mathcal{K}$-factors are presented in the lower panels for each contribution separately alongside their respective scale uncertainties. Uncertainties related to the choice of μ_{EFT} are not shown. All results are evaluated using the NNPDF3.1 PDF set and $\mu_R = \mu_F = H_T/2$ as well as $\mu_{EFT} = m_t$.*

σ_{tG} and $\sigma_{\phi G,\phi G}$, where both observables see effects of up to 25% in the regions where $\Delta\phi_{\ell_1,\ell_2}/\pi = 0$ or $\Delta\phi_{\ell_1,\ell_2}/\pi = 1$. One interesting observation here is that this shape difference to the SM is not observed in the squared contribution $\sigma_{tG,tG}$. This can be likely attributed to the chirality-structure of the gluon-top vertex. The $\mathcal{O}_{tG}$ operator induces terms of the type $\sim \left(\bar{Q}\,\sigma^{\mu\nu}\tilde{\phi}\,t\right) G^a_{\mu\nu}$, which flip the chiralities relative to the SM vertex $\sim \bar{Q}\gamma^\mu Q G^a_\mu$. The interference of these structures leads to the difference in the spin-correlation sensitive observables. For the squared contributions this effect consequently does not appear anymore, because the chirality-flip is present in both terms of the squared matrix element. The squared contribution $\sigma_{\phi G,\phi G}$ sees large shape changes in the range of 20% to 40% for $\Delta R(\ell_1,\ell_2)$ and between 30% and 65% for $\Delta\phi_{\ell_1,\ell_2}$. On the other hand, the linear contribution $\sigma_{\phi G}$ shows only very slight deviation from the SM shape.

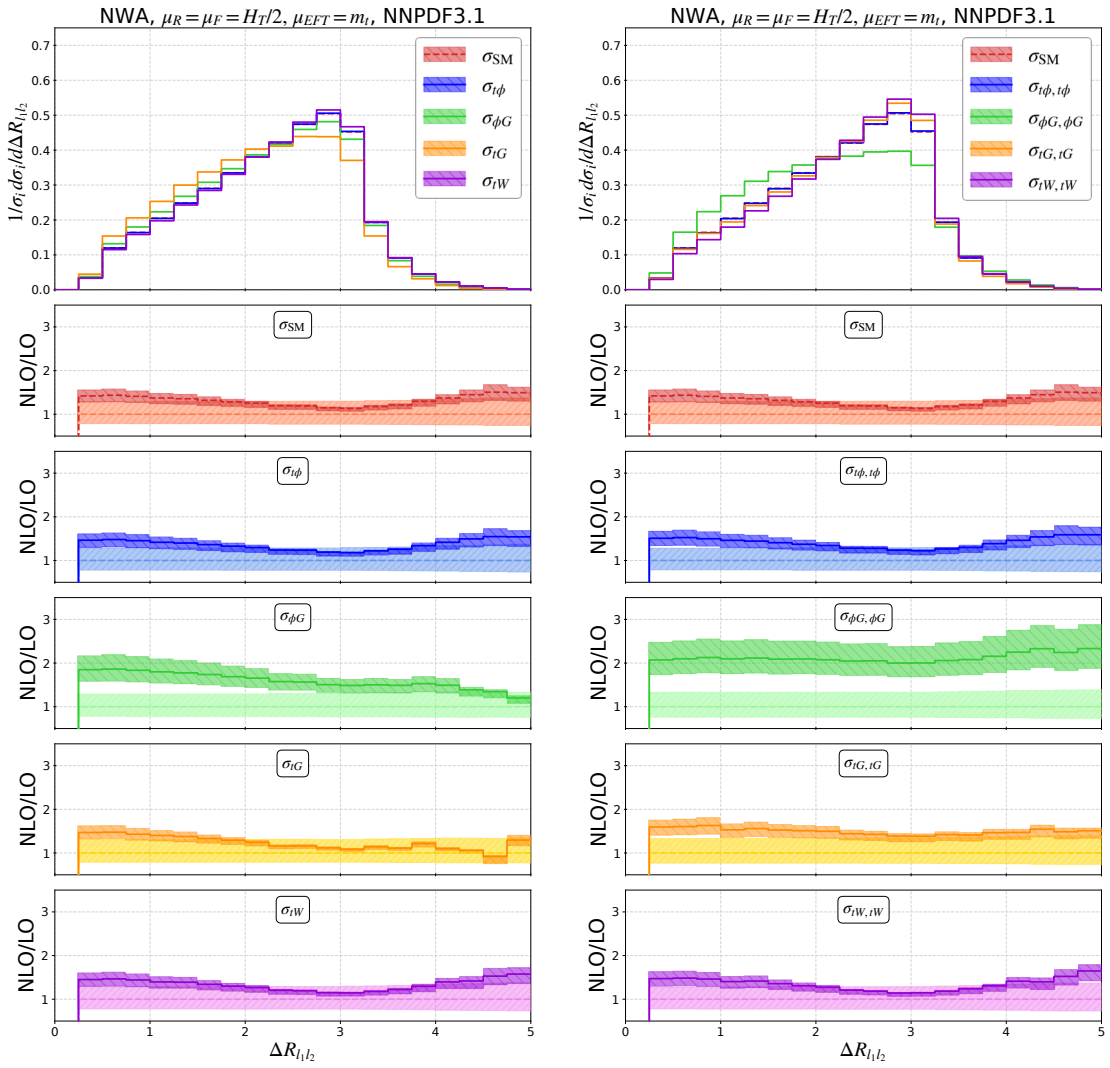


Figure 6.14: *Normalized differential cross section distributions in $\Delta R_{\ell_1\ell_2}$ for the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process at NLO QCD. The first panel shows the SM predictions as well as the linear (squared) SMEFT contributions in the left (right) column. Differential $\mathcal{K}$-factors are presented in the lower panels for each contribution separately alongside their respective scale uncertainties. Uncertainties related to the choice of μ_{EFT} are not shown. All results are evaluated using the NNPDF3.1 PDF set and $\mu_R = \mu_F = H_T/2$ as well as $\mu_{EFT} = m_t$.*

We briefly return to the topic of the large $\mathcal{K}$-factors in the $\mathcal{O}_{\phi G}$ partial contributions. Studying the impact of a jet-veto on the differential level allows us to further demonstrate the impact of the LO-precise diagrams on the calculation. To that end, we present the y_H distribution in Figure 6.16 and the azimuthal angle between the leptons $\Delta\phi_{\ell_1,\ell_2}$ in Figure 6.17. In this plot, we display the NLO shape after applying a jet-veto of 100 GeV and compare it to the corresponding LO and NLO predictions. The left plot shows the $\sigma_{\phi G}$ term and the right one the $\sigma_{\phi G,\phi G}$. In the lower panels, we present the corresponding scale uncertainties for all three configurations.

We first note that the shape of these distribution does not change when applying the veto, which indicates that the weight of the real-emission diagrams are distributed equally throughout the histograms. The theoretical uncertainty of the

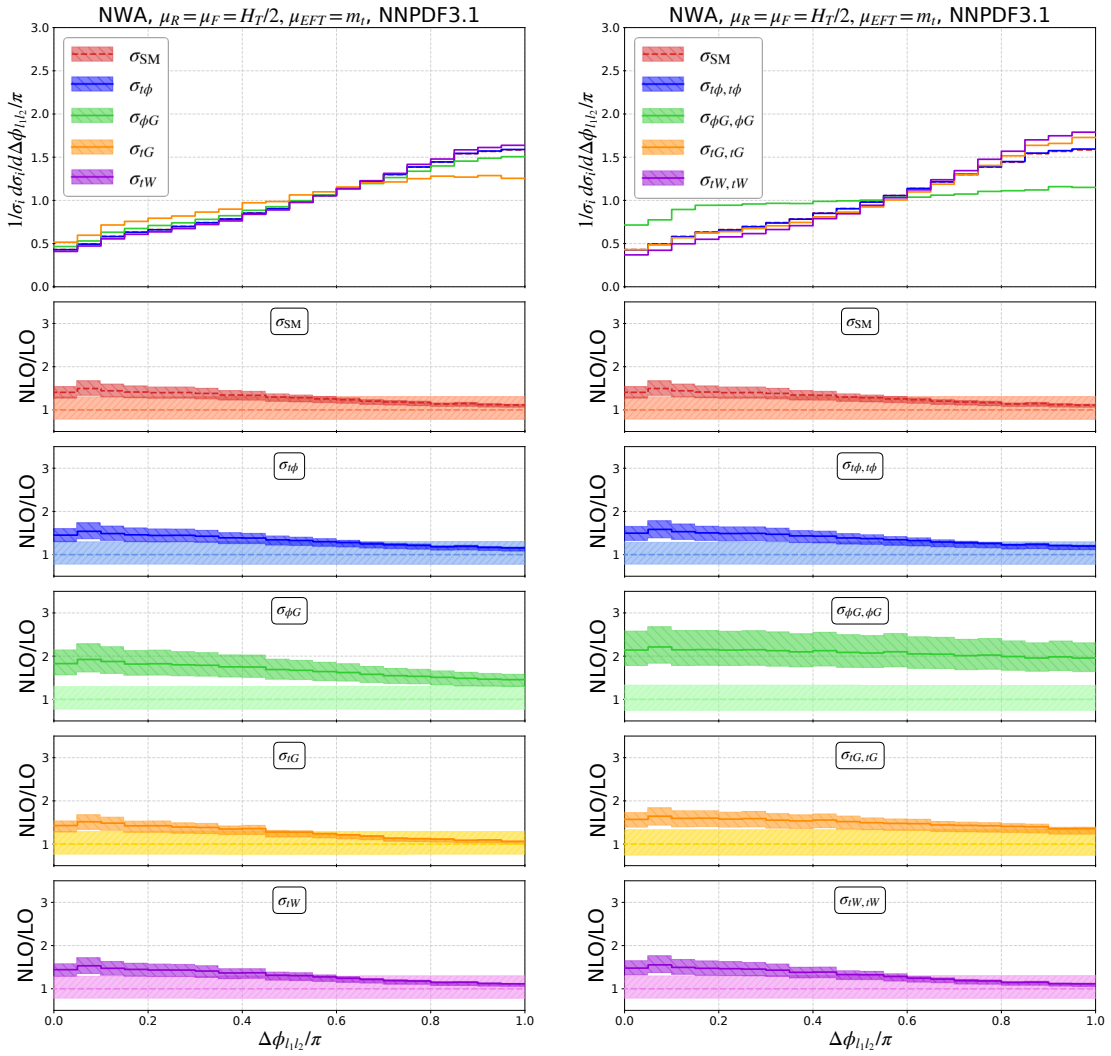


Figure 6.15: *Normalized differential cross section distributions in $\Delta\phi_{\ell_1\ell_2}/\pi$ for the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process at NLO QCD. The first panel shows the SM predictions as well as the linear (squared) SMEFT contributions in the left (right) column. Differential $\mathcal{K}$-factors are presented in the lower panels for each contribution separately alongside their respective scale uncertainties. Uncertainties related to the choice of μ_{EFT} are not shown. All results are evaluated using the NNPDF3.1 PDF set and $\mu_R = \mu_F = H_T/2$ as well as $\mu_{EFT} = m_t$.*

LO terms are at the level of $30\% - 40\%$, but get reduced at NLO to around 10% for linear and 20% for the quadratic term. Employing the jet-veto leads to large reduction to about 5% (10%) for y_H for linear (quadratic) contributions. Similarly, for $\Delta\phi_{\ell_1,\ell_2}$ we find a reduction to about $10\% - 15\%$.

A similar picture emerges for dimensionful observables. We present the transverse momentum of the Higgs boson and of the $b\bar{b}$ system in Figures 6.18 and 6.19, respectively. We note that, unlike the angular distributions, these observables are not well behaved when employing a fixed scale setting even with only a moderate p_T^{veto}. The dynamical scale choice is needed to provide sufficient perturbative stability. For the former, the uncertainty is reduced from up to 15% to about 8% in the linear contribution. For the squared one, the uncertainty in the low $p_T(H)$ region

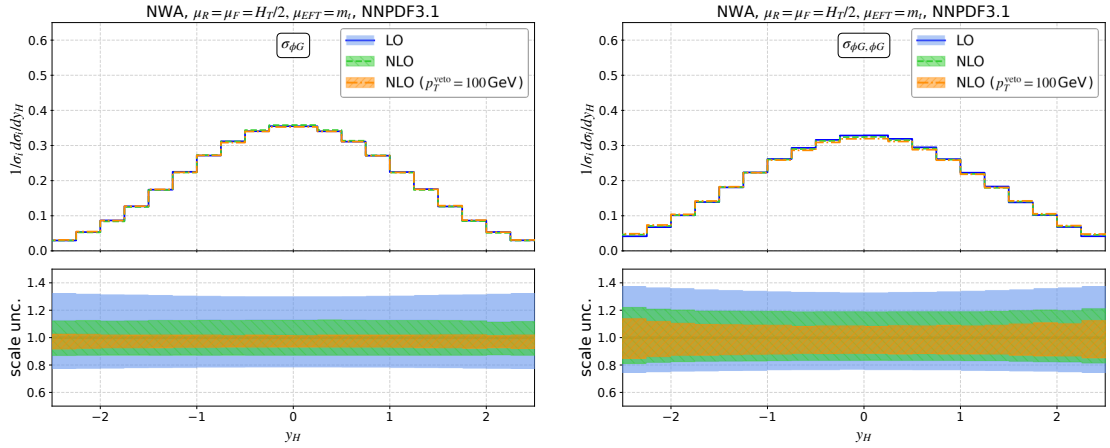


Figure 6.16: *Comparison of normalized differential y_H distribution in the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process. The left and right panel present the linear and squared $\mathcal{O}_{\phi G}$ contributions, respectively. Results are compared at LO, NLO and NLO with a jet-veto of $p_T^{veto} = 100\,GeV$. The lower panels compares the differential scale uncertainties obtained using 7-point scale variation. All results are evaluated using $\mu_R = \mu_F = H_T/2$ as well as $\mu_{EFT} = m_t$. The PDF set NNPDF3.1 is employed in all cases. Taken from Ref. [3].*

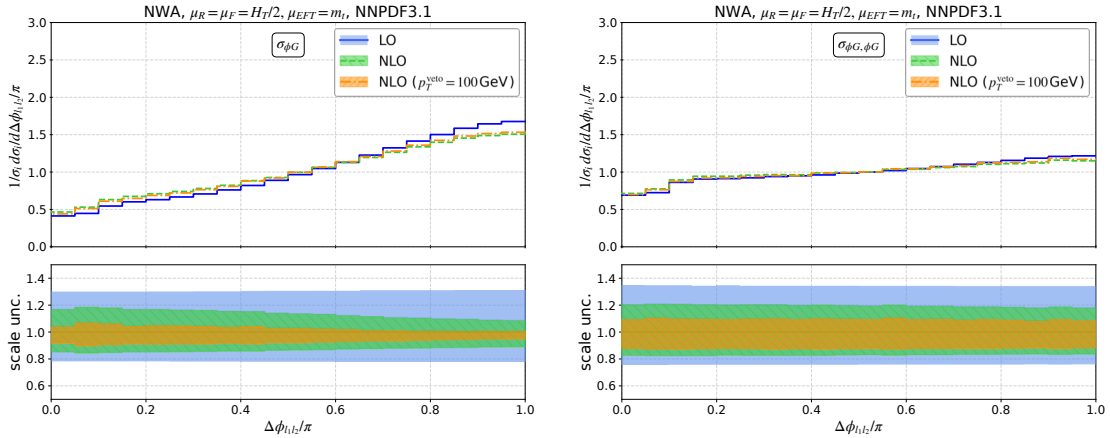


Figure 6.17: *Comparison of normalized differential $\Delta\phi_{\ell_1\ell_2}/\pi$ distribution in the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process. The left and right panel present the linear and squared $\mathcal{O}_{\phi G}$ contributions, respectively. Results are compared at LO, NLO and NLO with a jet-veto of $p_T^{veto} = 100\,GeV$. The lower panels compares the differential scale uncertainties obtained using 7-point scale variation. All results are evaluated using $\mu_R = \mu_F = H_T/2$ as well as $\mu_{EFT} = m_t$. The PDF set NNPDF3.1 is employed in all cases. Taken from Ref. [3].*

is about 20% at NLO and gets reduced 15% with the jet-veto, while the tail of the distribution sees a reduction from 15% to 7%. Looking at the $p_T(b_1b_2)$ histogram we similarly find that the uncertainties are reduced up to 10 percentage points in some regions. Overall we can conclude that the aforementioned LO-like real-emission diagrams drive the large uncertainty for predictions with the $\mathcal{O}_{\phi G}$ operator at NLO.

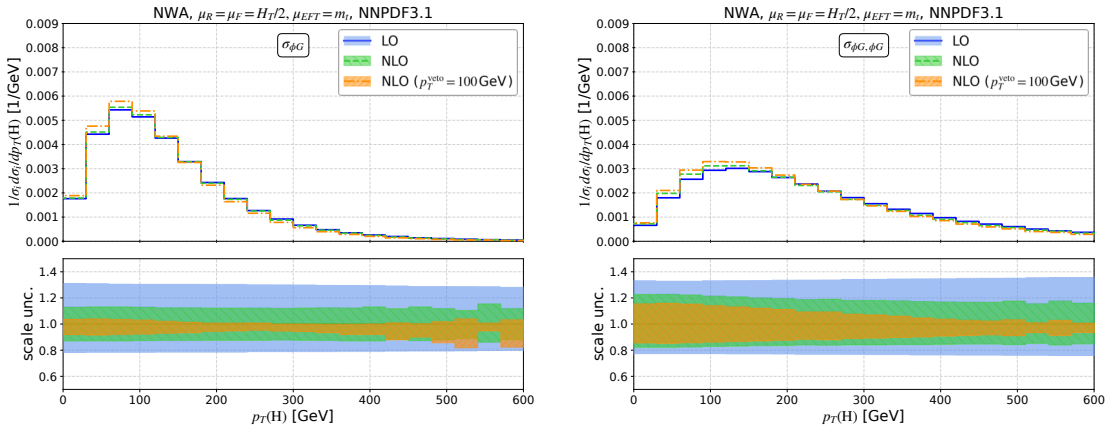


Figure 6.18: *Comparison of normalized differential $p_T(H)$ distribution in the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process. The left and right panel present the linear and squared $\mathcal{O}_{\phi G}$ contributions, respectively. Results are compared at LO, NLO and NLO with a jet-veto of $p_T^{veto} = 100\,GeV$. The lower panels compares the differential scale uncertainties obtained using 7-point scale variation. All results are evaluated using $\mu_R = \mu_F = H_T/2$ as well as $\mu_{EFT} = m_t$. The PDF set NNPDF3.1 is employed in all cases.*

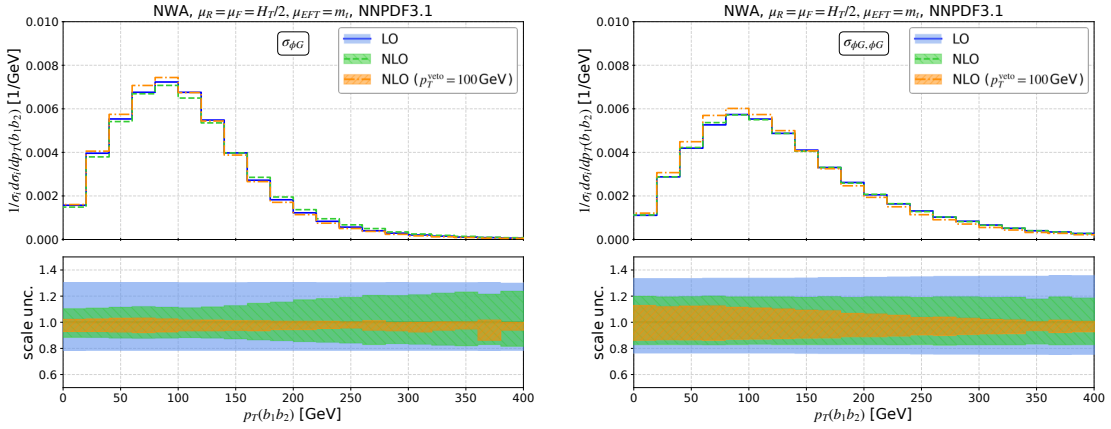


Figure 6.19: *Comparison of normalized differential $p_T(b_1b_2)$ distribution in the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process. The left and right panel present the linear and squared $\mathcal{O}_{\phi G}$ contributions, respectively. Results are compared at LO, NLO and NLO with a jet-veto of $p_T^{veto} = 100\,GeV$. The lower panels compares the differential scale uncertainties obtained using 7-point scale variation. All results are evaluated using $\mu_R = \mu_F = H_T/2$ as well as $\mu_{EFT} = m_t$. The PDF set NNPDF3.1 is employed in all cases.*

6.4 Fixed and dynamical scale choice comparison

The usage of dynamical, phase-space dependent, choices for the renormalization and factorization scales is often preferable over fixed definitions for the calculation of differential distributions. While fixed choices usually work very well in inclusive observables, one typically finds distorted $\mathcal{K}$-factors in the more exclusive regimes of the phase-space, where many differential observables are defined. So far we have discussed differential observables evaluated with the dynamical choice $\mu_R = \mu_F =$

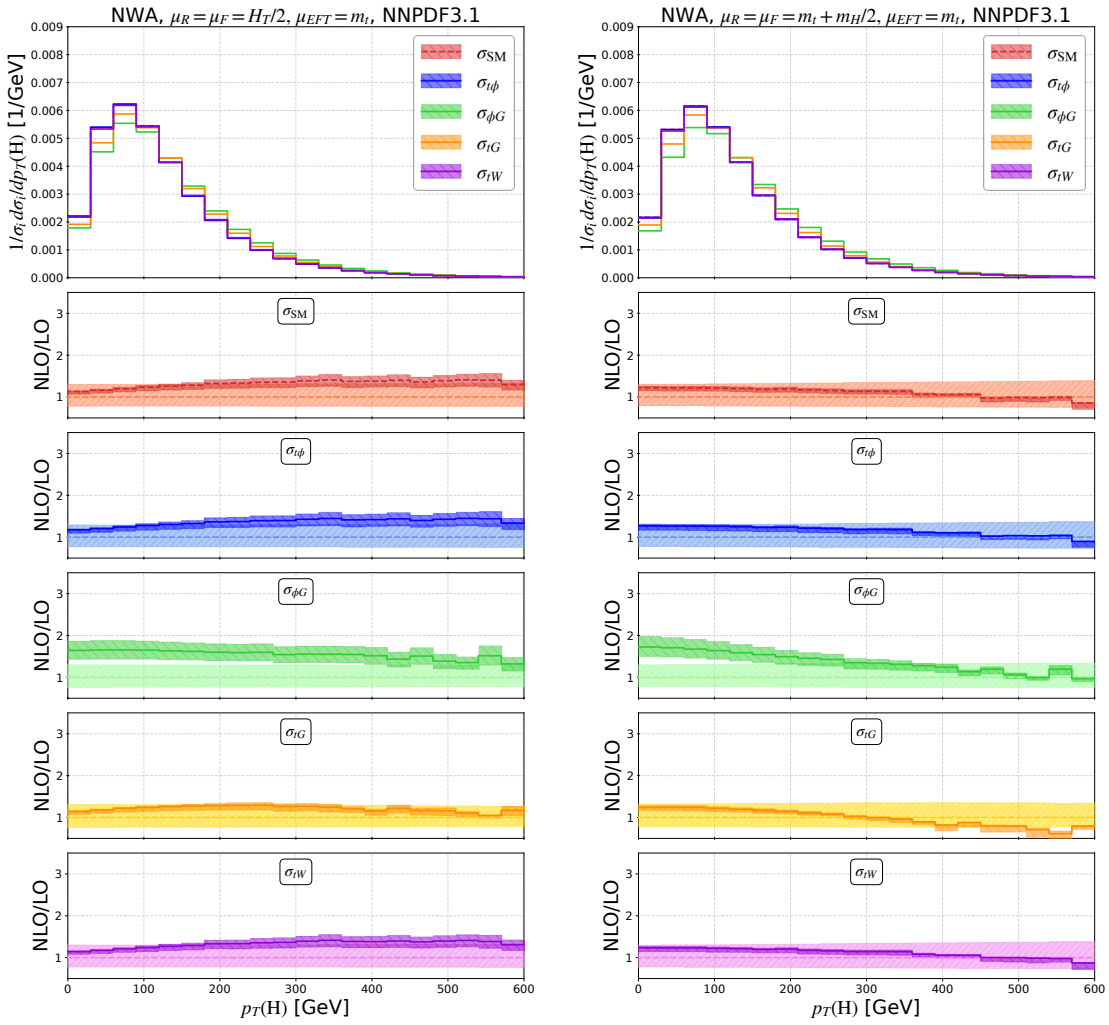


Figure 6.20: *Normalized differential cross section distribution in $p_T(H)$ for the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process at NLO QCD. Lower panels display differential $\mathcal{K}$-factors alongside their scale uncertainty. The left column uses the dynamical scale setup $\mu_R = \mu_F = H_T/2$, while the right one uses $\mu_R = \mu_F = m_t + m_H/2$. The Wilson coefficients are defined at $\mu_{EFT} = m_t$ in both cases. Only linear SMEFT contributions alongside the SM prediction is shown. The NNPDF3.1 PDF set is used to evaluate the histograms.*

$H_T/2$. We will now briefly investigate the differences and similarities when a fixed value of $\mu_R = \mu_F = m_t + m_H/2$ is used instead. We always set $\mu_{EFT} = m_t$ to retain comparability. As before all results are evaluated using the NNPDF3.1 PDF set.

We start with the Higgs kinematics. The transverse momentum of the Higgs, $p_T(H)$, is shown in Figure 6.20. We only show the linear SMEFT contributions as well as the SM predictions here. The left column is computed using the dynamical scale choice, while the right column uses the fixed one. We observe very similar shapes of the NLO distributions as can be seen in the first row. Differences arise primarily in the differential $\mathcal{K}$-factors for this observable. We find that both SMEFT and SM terms are strongly angled when employing the fixed scale, while the dynamical scale generates much flatter NLO QCD corrections. This points towards better perturbative convergence for the dynamical scale choice, which could

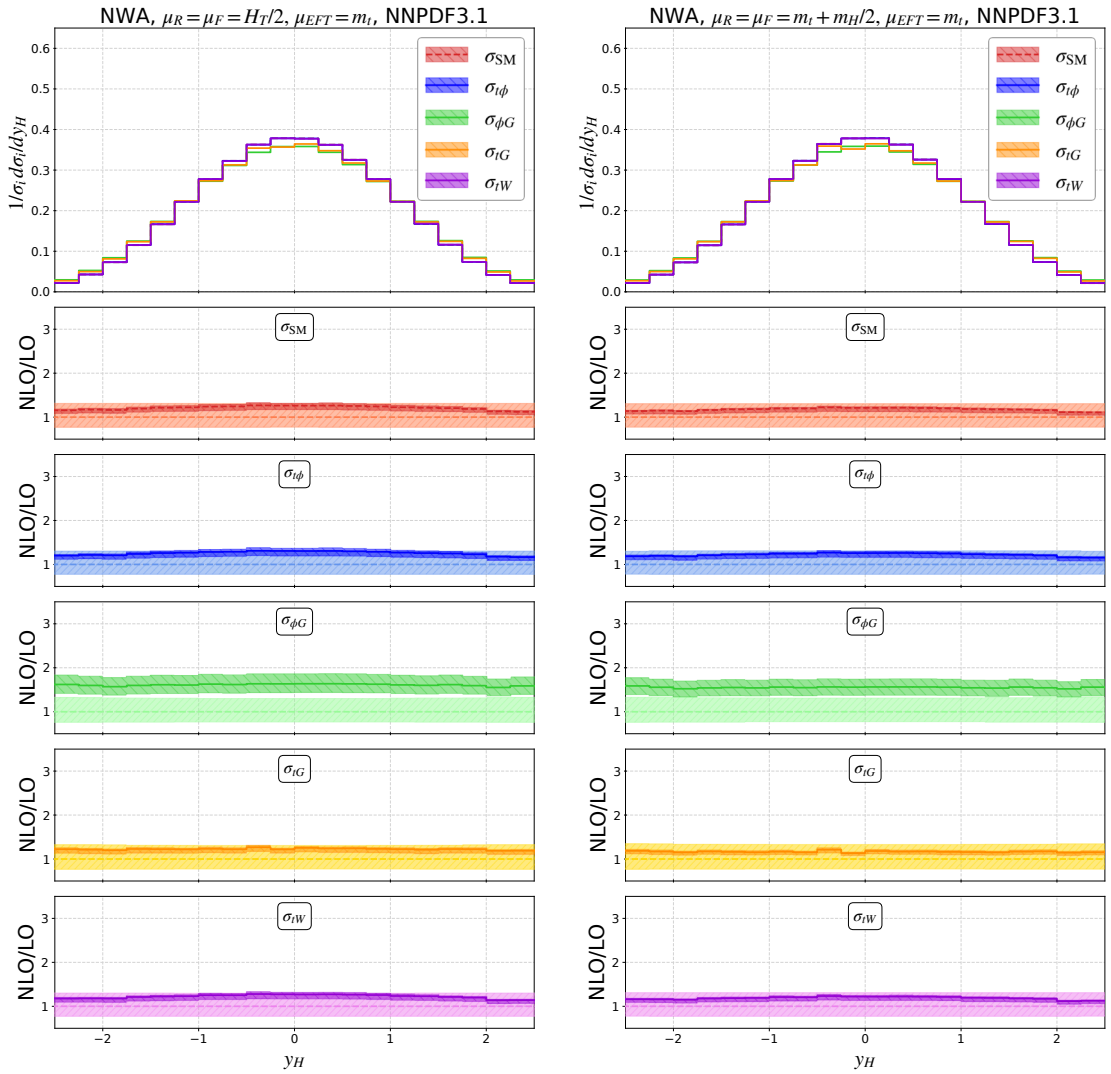


Figure 6.21: *Normalized differential cross section distribution in y_H for the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process at NLO QCD. Lower panels display differential $\mathcal{K}$-factors alongside their scale uncertainty. The left column uses the dynamical scale setup $\mu_R = \mu_F = H_T/2$, while the right one uses $\mu_R = \mu_F = m_t + m_H/2$. The Wilson coefficients are defined at $\mu_{EFT} = m_t$ in both cases. Only linear SMEFT contributions alongside the SM prediction is shown. The NNPDF3.1 PDF set is used to evaluate the histograms.*

impact the normalization of bins in specific phase-space regions. While agreement with the LO scale bands is arguably slightly better in the fixed scale case, we stress that the choice of $\alpha_s(m_Z)$ used in the LO calculations has a strong impact on these conclusions. Choosing $\alpha_s(m_Z) = 0.130$ instead of $\alpha_s(m_Z) = 0.118$ would reduce the $\mathcal{K}$-factors by about 19%. The size of scale uncertainties is comparable for both scale choices [1].

Dimensionless observables like angular distribution and rapidities typically aren't sensitive to the scale choice to the same degree. We provide the example of the Higgs rapidity distribution y_H in Figure 6.21. We again focus on linear contributions. In

[1] Note that the size of the NLO scale uncertainty is computed relative to the central value of the NLO band. One can not directly read off the result from the provided figure.

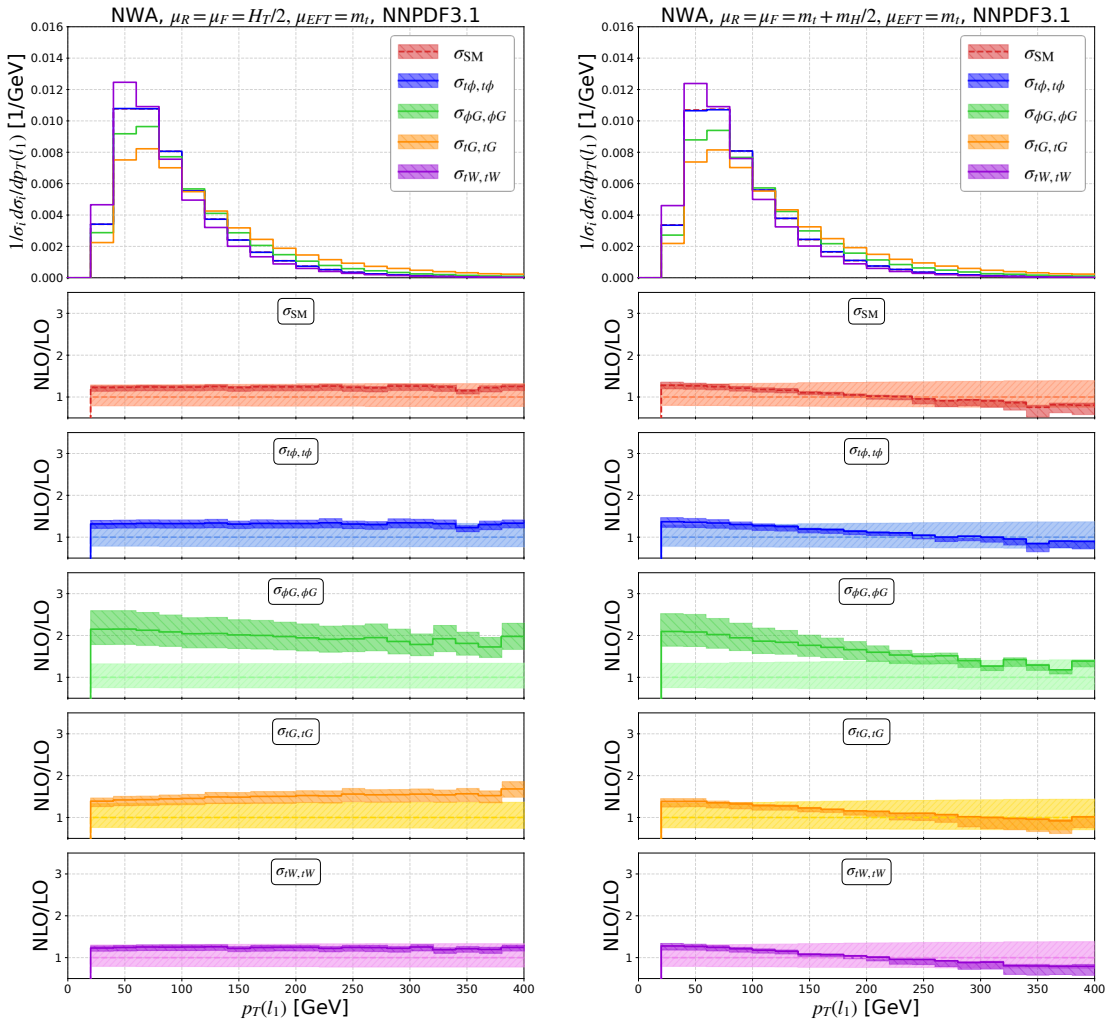


Figure 6.22: *Normalized differential cross section distribution in $p_T(l_1)$ for the $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ process at NLO QCD. Lower panels display differential $\mathcal{K}$-factors alongside their scale uncertainty. The left column uses the dynamical scale setup $\mu_R = \mu_F = H_T/2$, while the right one uses $\mu_R = \mu_F = m_t + m_H/2$. The Wilson coefficients are defined at $\mu_{EFT} = m_t$ in both cases. Only squared SMEFT contributions alongside the SM prediction is shown. The NNPDF3.1 PDF set is used to evaluate the histograms.*

this case both the shape of the distributions well as the differential $\mathcal{K}$-factors match closely. Such observables are generally well described with both scale choices.

We conclude by studying one exemplary top-quark decay product. Figure 6.22 shows the transverse momentum of the hardest lepton $p_T(l_1)$. While we focused on the linear terms so far, we find similar conclusions also for quadratic terms, which we demonstrate with this observable. We can again observe that the shape of the distributions is very similar again. However, the size of the NLO QCD corrections is significantly flatter for the dynamical scale choice than the fixed one. For this observable, the differential $\mathcal{K}$-factor vary by more than the entire size of the LO uncertainty band throughout this distribution. This observation is common for both the SMEFT and SM predictions.

We again conclude that the use of dynamical scale choices has little impact

on the shape of the distributions studied in this section. However, we find much better perturbative convergence when investigating the unnormalized distributions for dimensionful distribution, which manifests in much flatter differential $\mathcal{K}$-factors. In particular, high p_T regimes are affected, which are relevant for many BSM studies.

6.5 Comparison with a stable top-quark computation

To determine the impact of fiducial phase-space cuts and top-quark decay modelling on the predictions in this process, we present a comparison of several observables for $pp \to t\bar{t}H+X$ and $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H+X$ in this section. Naturally, the discussion is limited to observables related to the production process only. To perform this comparison, we must determine the top-quark kinematics from the decay products in the NWA case. The reconstruction of the top-quark momenta in the NWA comes with one subtlety. b-quark initiated real-emission processes can include up to three b-jets in their final state. Furthermore, one of the b-jets originating from the top-quark decays can be unresolved due to the phase-space selection. To consistently define the top-quark momenta in the NWA, we thus have to define a reconstruction procedure. To that end, we evaluate the quantity

$$Q = |M(W^+b_1) - m_t| + |M(W^-b_2) - m_t| \tag{6.21}$$

for all pairs of b-jets (b_i, b_j) in the event and choose the combination that minimizes the value of Q. We assume perfect reconstruction of the W boson kinematics for this analysis. We tested explicitly that including the light-jet in the definition of Q does not lead to any noticeable difference in the kinematic distributions. We further note that in the NWA this reconstruction is expected to be extremely accurate due to the on-shell requirement of the top-quarks in the calculation. To make the two computational setups more comparable we opt to present results using the fixed scale setup in both top-quark modelling approaches, that is we set $\mu_R = \mu_F = m_t + m_H/2$ and $\mu_{EFT} = m_t$. The dynamical scale setup would naturally introduce some mismatch as the corresponding final states differ.

Among the top-quark observables we focus on the transverse momentum of one of the top-quarks $p_T(t)$ as well as the sum of transverse momenta of the two tops and the Higgs boson defined as

$$H_T^{reco.} = p_T(t) + p_T(\bar{t}) + p_T(H)\,. \tag{6.22}$$

The kinematics of the Higgs boson are also interesting as they do not depend on the modelling approach employed for the top-quarks. The two representative distributions in this case are the transverse momentum $p_T(H)$ and the rapidity y_H.

The figures are organized as follows. In the upper panel we plot the ratio of the normalized NWA prediction to normalized prediction using stable top-quarks. This quantity effectively measures the relative shape difference of the computational approaches. A value of 1 means that the two calculations give equal predictions of the shape of the distributions, while a slope characterizes shape differences. We plot this ratio for each operator separately and also provide the SM result for comparison. As before, the left column corresponds to the linear terms, while the right column

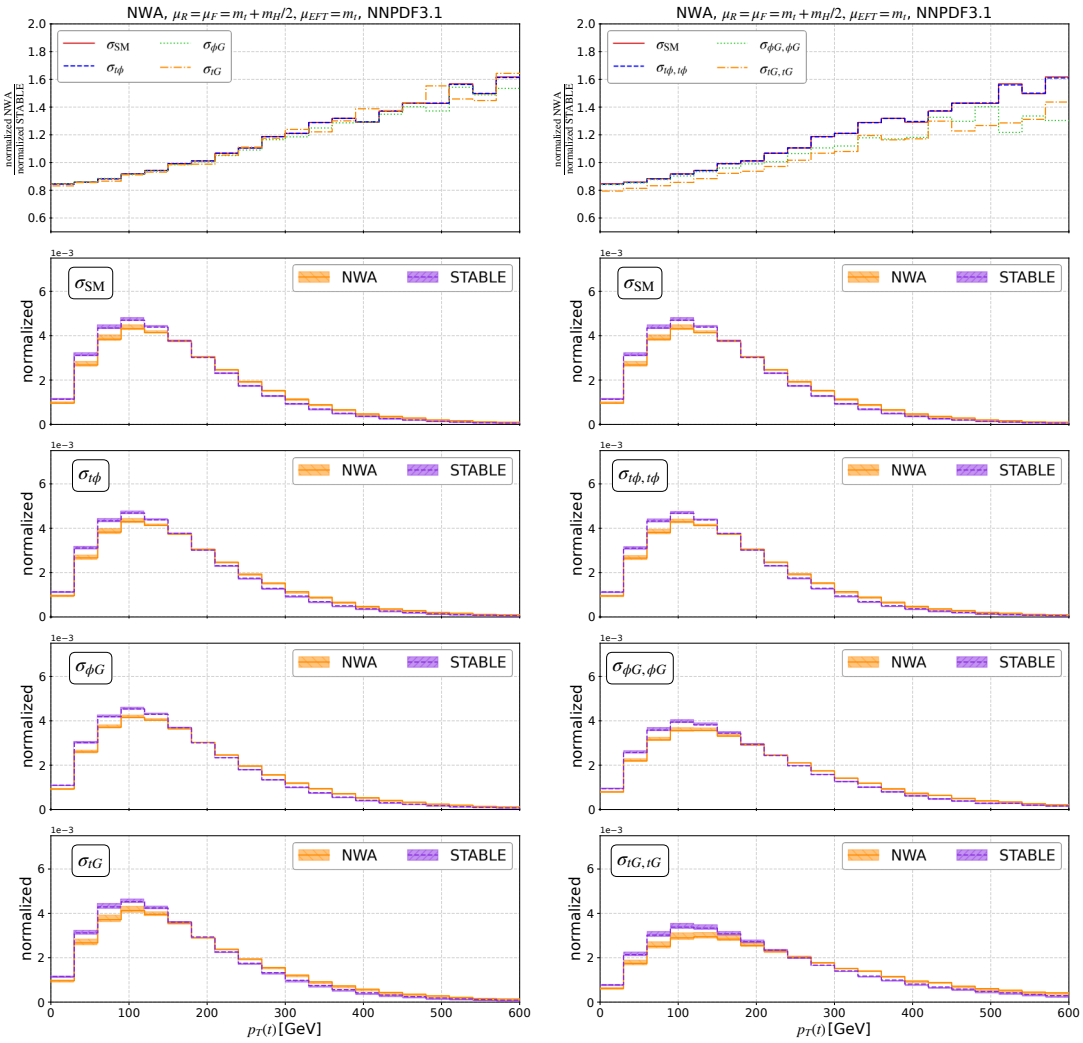


Figure 6.23: *Ratio of normalized* $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ *and* $pp \to t\bar{t}H + X$ *differential cross section distribution as a function of* $p_T(t)$*. The left column presents linear SMEFT contributions, while the right one shows the squared terms. The SM is also displayed for comparison. The lower panels compare each curve individually together with their respective scale uncertainties. The results are evaluated using* $\mu_R = \mu_F = m_t + m_H/2$ *and* $\mu_{EFT} = m_t$ *as well as the NNPDF3.1 PDF set. The center-of-mass energy is set to* $\sqrt{s} = 13.6$ *TeV. Taken from Ref. [3].*

contains squared contributions. The interference terms are not plotted to increase the clarity of the results. The lower rows contain a comparison of the normalized cross section for the two computations for each operator individually. The scale uncertainties based on the 7-point variation of μ_R and μ_F are also shown.

The transverse momentum is presented in Figure 6.23. The linear terms σ_i behave very similarly to the SM here. In both cases, we observe shape distortions of about 15% at low values of $p_T(t)$ up to 65% at the tail of the distribution. Compared to the theoretical uncertainty of about 10% for all terms throughout most of the spectrum this represents a large deviation, which indicates the importance of including a fiducial phase-space selection. The quadratic case sees pronounced differences for $\mathcal{O}_{tG}$ and $\mathcal{O}_{\phi G}$ compared to the SM and $\mathcal{O}_{t\phi}$. The inclusion of these SMEFT operators in the matrix element yields a partial cancellation against the

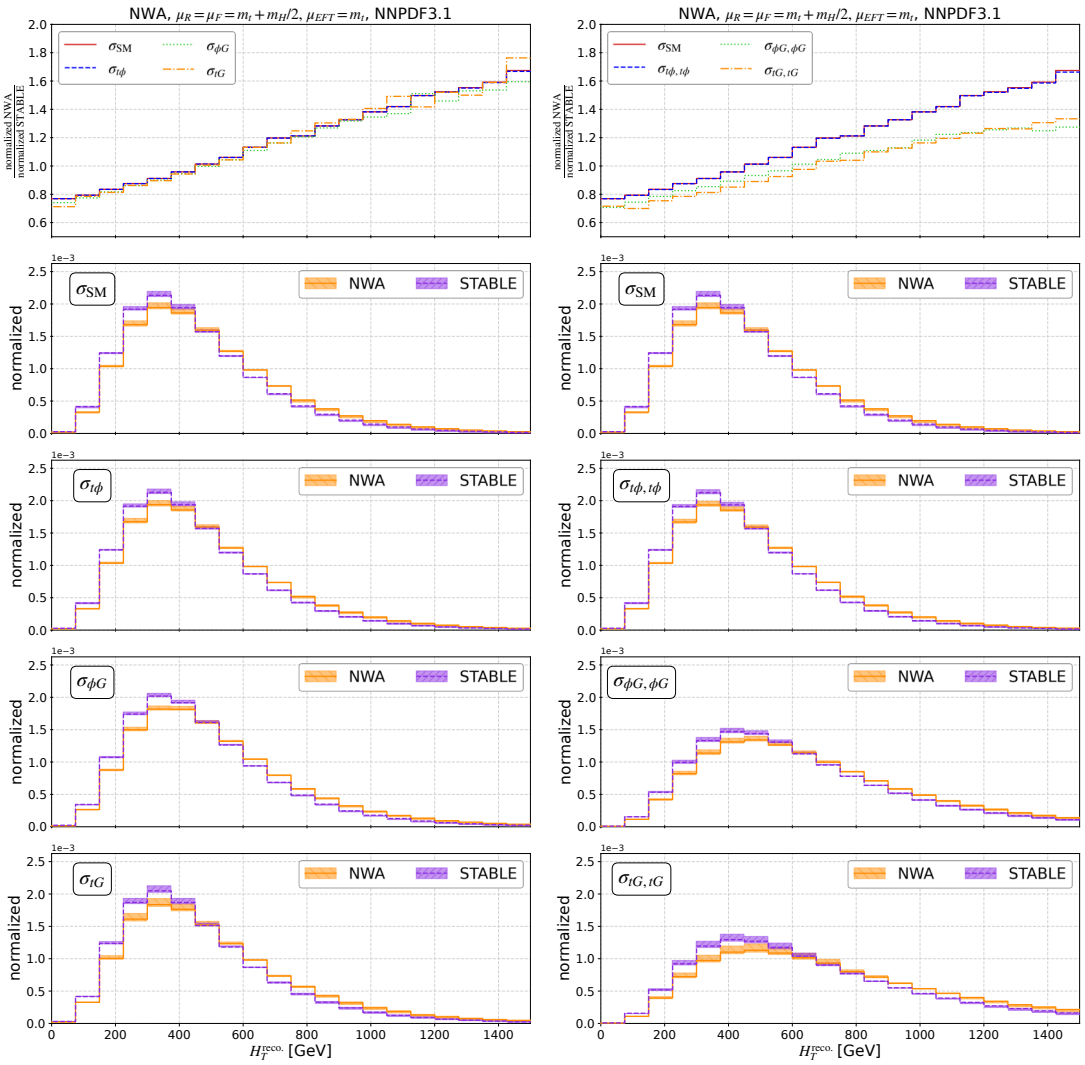


Figure 6.24: *Ratio of normalized* $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ *and* $pp \to t\bar{t}H + X$ *differential cross section distribution as a function of* H_T^{reco}*. The left column presents linear SMEFT contributions, while the right one shows the squared terms. The SM is also displayed for comparison. The lower panels compare each curve individually together with their respective scale uncertainties. The results are evaluated using* $\mu_R = \mu_F = m_t + m_H/2$ *and* $\mu_{EFT} = m_t$ *as well as the NNPDF3.1 PDF set. The center-of-mass energy is set to* $\sqrt{s} = 13.6$ *TeV. Taken from Ref. [3].*

kinematic effects present in the SM. As a result the differences are slightly increased for softer top-quarks, reaching about 20%, but are decreased to about 30% to 40% for the hard part of the spectrum.

The equivalent distribution for H_T^{reco} is shown in Figure 6.24, which provides an overall similar conclusion to $p_T(t)$. The scale uncertainties are between 5% and 15%. We observe shape distortions of about 90% for both the SM and the linear partial contributions. The quadratic terms $\sigma_{tG,tG}$ and $\sigma_{\phi G,\phi G}$ see a partial cancellation of these kinematic effects against the SMEFT induced effects. In that case we find that the tail of the distribution rises at most to about 30% compared to the 60% for the SM.

Moving on to the Higgs boson kinematics, we first focus on the transverse mo-

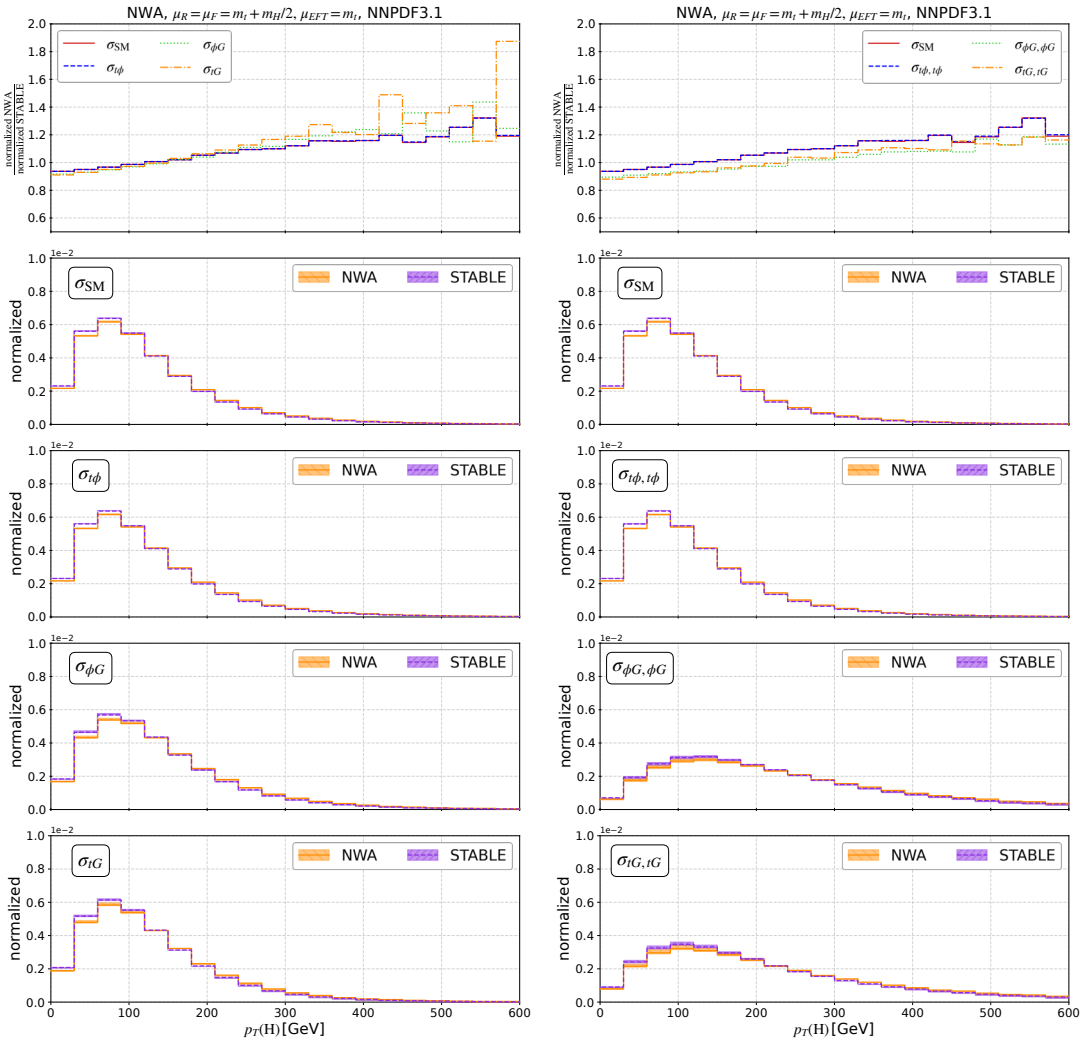


Figure 6.25: *Ratio of normalized* $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ *and* $pp \to t\bar{t}H + X$ *differential cross section distribution as a function of* $p_T(H)$*. The left column presents linear SMEFT contributions, while the right one shows the squared terms. The SM is also displayed for comparison. The lower panels compare each curve individually together with their respective scale uncertainties. The results are evaluated using* $\mu_R = \mu_F = m_t + m_H/2$ *and* $\mu_{EFT} = m_t$ *as well as the NNPDF3.1 PDF set. The center-of-mass energy is set to* $\sqrt{s} = 13.6$ *TeV. Taken from Ref. [3].*

mentum $p_T(H)$ shown in Figure 6.25 [2]. We first note that the differences in the kinematics of the Higgs are much smaller in the two computational approaches. At the largest points they reach about 25%. The normalized distributions carry rather small theoretical uncertainties of about 5% (10%) for the linear (squared) terms in the first half of the histogram. Consequently, the shape difference is largely not within the theoretical uncertainty estimate, which is not surprising given that scale-uncertainties only estimate missing higher order contributions in the perturbative expansion. For quadratic terms there is again a small difference for $\mathcal{O}_{tG}$ and $\mathcal{O}_{\phi G}$ compared to the SM, but the effects are much less pronounced than for the top-quark

[2] We note that the statistical accuracy of the Monte Carlo integration is low in the tail of this distribution for the σ_{tG} and $\sigma_{\phi G}$ cases only. This behavior is driven by the very low event count of both numerator and denominator in this region of the phase-space.

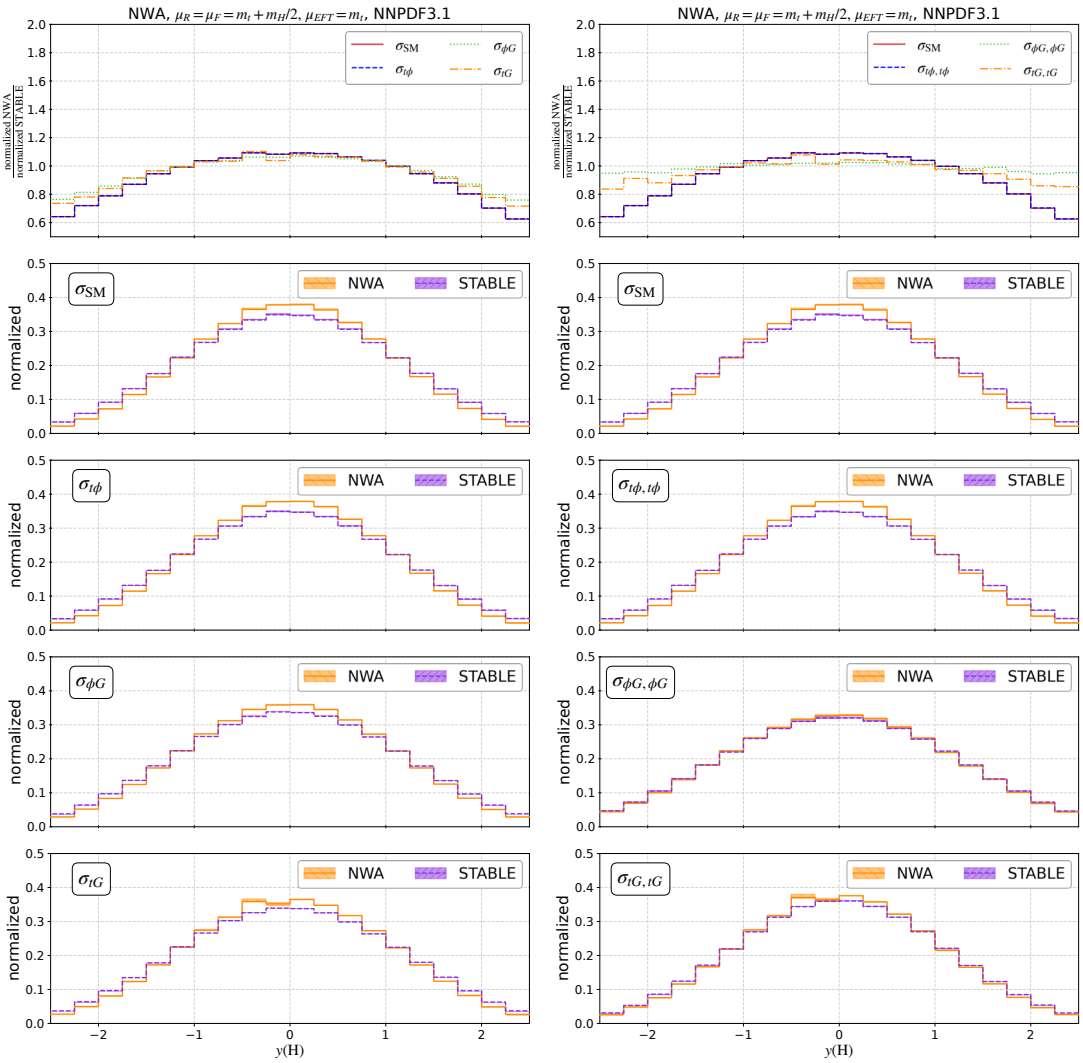


Figure 6.26: *Ratio of normalized* $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ *and* $pp \to t\bar{t}H + X$ *differential cross section distribution as a function of* y_H*. The left column presents linear SMEFT contributions, while the right one shows the squared terms. The SM is also displayed for comparison. The lower panels compare each curve individually together with their respective scale uncertainties. The results are evaluated using* $\mu_R = \mu_F = m_t + m_H/2$ *and* $\mu_{EFT} = m_t$ *as well as the NNPDF3.1 PDF set. The center-of-mass energy is set to* $\sqrt{s} = 13.6$ *TeV. Taken from Ref. [3].*

kinematics.

The rapidity distribution of the Higgs boson y_H is presented in Figure 6.26. Differences are larger compared to $p_T(H)$. In particular, towards the edges of the spectrum ($y_H \approx 2.5$) one can observe a ratio of about 35%. Notable is that the σ_{tG} and $\sigma_{\phi G}$ contributions already show a partial cancellation of these effects, where the difference to 1 stays below the 25% level. This is further amplified in the quadratic terms $\sigma_{tG,tG}$ and $\sigma_{\phi G,\phi G}$. In the former case the two computations agree within 20%, while the latter is identical for practical reasons.

Lastly, the azimuthal angle between the two top-quarks is shown in Figure 6.27. This angular distribution is important in the context of spin-correlation measurements, which could be modified in many common BSM scenarios. The impact of

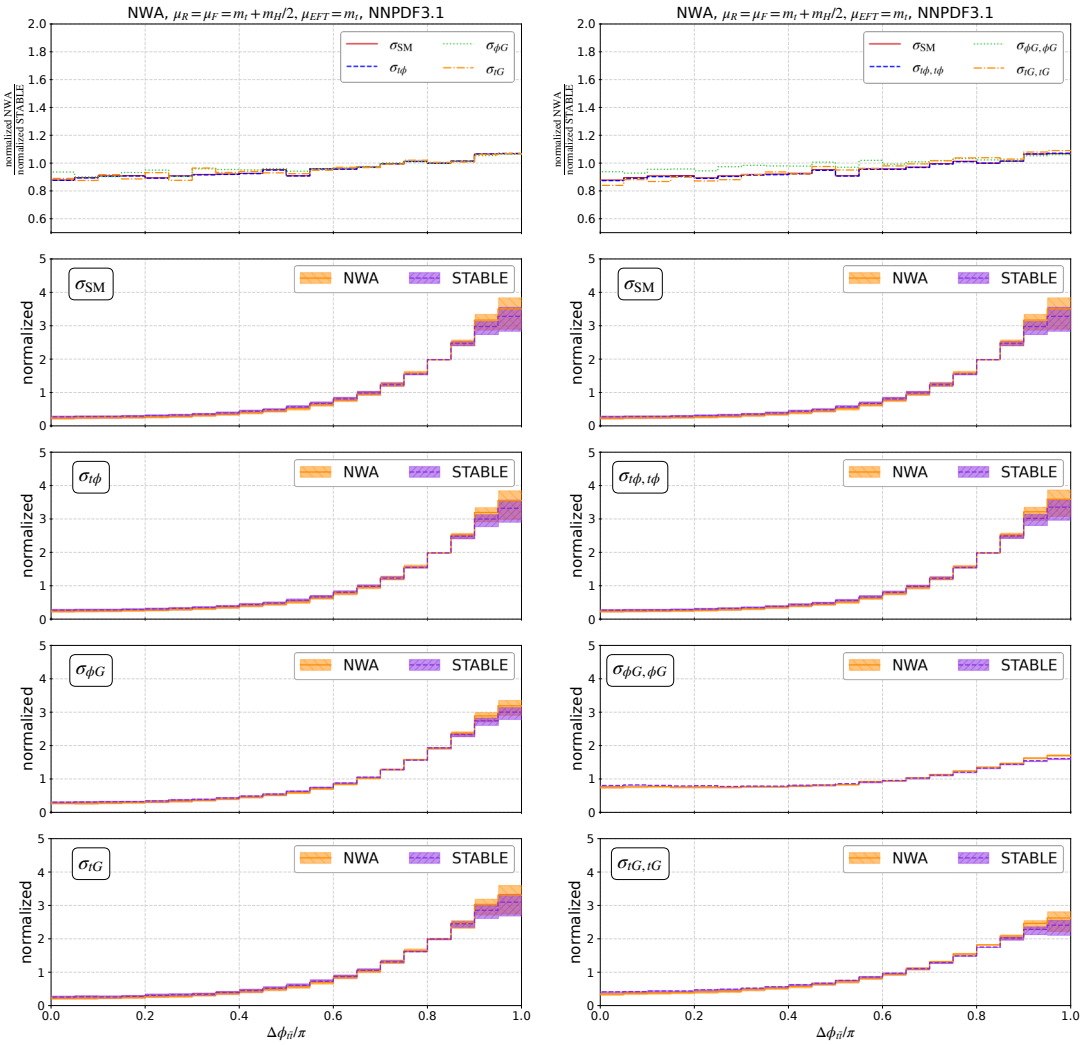


Figure 6.27: *Ratio of normalized $pp \to e^+\nu_e\,\mu^-\bar{\nu}_\mu\,b\bar{b}\,H + X$ and $pp \to t\bar{t}H + X$ differential cross section distribution as a function of $\Delta\phi_{t\bar{t}}/\pi$. The left column presents linear SMEFT contributions, while the right one shows the squared terms. The SM is also displayed for comparison. The lower panels compare each curve individually together with their respective scale uncertainties. The results are evaluated using $\mu_R = \mu_F = m_t + m_H/2$ and $\mu_{EFT} = m_t$ as well as the NNPDF3.1 PDF set. The center-of-mass energy is set to $\sqrt{s} = 13.6$ TeV.*

phase-space cuts is generally less pronounced than for the other observables presented in this section. At the level of the linear contributions, we find similar behavior for the SM and the several SMEFT contributions. A shape differences of about 10% can be observed at both the beginning and end of the spectrum for our selection of four operators. Notably, this effect is roughly halved for the $\sigma_{\phi G,\phi G}$ term, while other quadratic terms behave like the SM. These are covered by the respective theoretical uncertainties as can be seen in the lower panels of the figure.

In conclusion, observed shape differences between calculations in the NWA and with stable top-quarks often do not extrapolate from the SM to all SMEFT contributions. Consequently, dedicated treatment in the unfolding can be required to properly treat effects of this kind in experimental analyses. Generally, even in the

SM we find that the very inclusive phase-space cuts used in our study can lead to large differences in kinematical distributions of the production process. These conclusions hold equally for dimensionless and dimensionful observables and the structure of corrections is generally non-trivial with both the observable definition and included SMEFT effects having a notable impact on the shape distortions.

Chapter 7

Summary and Conclusions

In this thesis, we presented theoretical predictions for the $pp \to t\bar{t}H + X$ process in the dileptonic decay channel using the NWA within in the SMEFT framework at NLO QCD using the LHC Run III center-of-mass energy $\sqrt{s} = 13.6\,\text{TeV}$. We studied the role of the $\mathcal{O}_{t\phi}$, $\mathcal{O}_{\phi G}$, $\mathcal{O}_{tG}$ and $\mathcal{O}_{tW}$ operators, which were consistently included in the production and decay parts of the process. Integrated results were presented for several scale choices. Differential results were analyzed for a variety of observables. We further examined the impact of a realistic final-state selection on the obtained differential distributions. To that end, we compared the results to a calculation with stable top quarks. To support both calculations, we developed and presented Helac-Smeft, an expanded version of the Helac-Nlo program. The new code is specialized in computing SMEFT contributions at NLO QCD and can include decays of unstable resonances using either the NWA or the full off-shell approach. The SM workflow was extended to the SMEFT use-case and fully supports the evaluation of theoretical uncertainties using an efficient reweighting procedure. Existing results from the literature can be interfaced via the UFO file format. This capability was utilized in this thesis with the Smeft@NLO UFO model. An significant number of additional optimizations is introduced to offset the additional computational cost of SMEFT calculations and we find that the overhauled program outperforms the original code in SM calculations as a result. On the integrated level, we find that the size of NLO QCD corrections can vary operator by operator. While most contributions see corrections of the order of $20 - 30\%$, we find that the $\mathcal{O}_{\phi G}$ operator typically sees much larger corrections. For terms linear in $C_{\phi G}$, we find that these reach around $60 - 70\%$ instead. This behavior is even more pronounced in the quadratic contribution, which sees an increase of more than 100% at NLO QCD. We identify the source of this behavior as additional topologies that open up within the real-emission channel, where the Higgs boson is emitted from the additional real radiation. We estimated two sources of theoretical uncertainties in this thesis. The first one is the typical scale variation related to the truncation of the perturbative expansion in α_s. The second one is related to the RGE running of the Wilson coefficients. Both are substantially reduced at NLO, although the exact magnitude depends on the operator under study. Excluding the $\mathcal{O}_{\phi G}$ operator, the uncertainty due to the α_s expansion is reduced by a factor two to three when going to NLO QCD. Due to the aforementioned new topologies at NLO, we find a much smaller reduction for the $\mathcal{O}_{\phi G}$ contributions. The theoretical error related to the defining scale of the Wilson coefficients is similarly reduced at NLO, although

the reduction of the uncertainty is less pronounced. We find that, in practice, the μ_R/μ_F variations are responsible for the majority of the theoretical uncertainty in our calculation. The situation is less clear when taking the ratio of SMEFT contributions to the SM prediction, where both types of uncertainties are similar in size and must both be taken into account. Comparing the stable calculation to the NWA, we find that uncertainties are slightly reduced in the latter. However, general trends are similar for the operators that are present in both calculations. We further find that in particular the ratio to the SM is well described in the stable top-quark calculation. Using a dedicated LO calculation with a fully inclusive final state selection, we explicitly demonstrate that the differences of the two computations can be attributed to the phase-space cuts imposed in the NWA computation. Moreover, we observe that the use of dynamical and fixed scales yields very similar results at the integrated level.

We studied kinematical features of the various SMEFT contributions compared to the SM within the NWA calculation. To that end, we focused on normalized distributions, which isolate the shape of the spectra. We find that the $\mathcal{O}_{\phi G}$ and $\mathcal{O}_{tG}$ operators generally yield harder p_T spectra compared to the SM and the other operators under investigation. This effect is considerably enhanced for squared dimension 6 terms. We further find that the angular distances of the $t\bar{t}$-pair are significantly altered for these two operators. The $\mathcal{O}_{t\phi}$ operator produces distributions which are usually very similar to the SM in shape, while the $\mathcal{O}_{tW}$ operator primarily modifies the shape of top-quark decay product kinematics. The leptons end up slightly softer, while the missing transverse momentum and p_T of the b-jet tend toward larger values compared to the SM. For the latter observable, one can only observe large effects for the squared contribution, while the impact on the leptonic observables is already very visible for the linear SMEFT contribution.

We extended the discussion of the additional LO-like topologies for the $\mathcal{O}_{\phi G}$ operator to the differential level. To that end, we imposed a jet-veto on the additional radiation and compared the size of the theoretical uncertainty to the regular calculation. We find that the uncertainties are substantially reduced as these topologies become more restricted, while the shape of observables stays largely intact. Together with the observations on the integrated level, we conclude that the aforementioned new topologies are driving the size of the theoretical uncertainty in NLO QCD calculations.

Next, we briefly discussed how a dynamical scale choice compares to a fixed one at the differential level. We observe that the dynamical scale choice generates much flatter differential $\mathcal{K}$-factors for dimensionful observables like $p_T(H)$, which points to more stable perturbative convergence. On the other hand, dimensionless quantities like the Higgs rapidity were largely unaffected by the scale definition. Comparing just the shapes of the distributions, we found that the predictions were similar for both setups.

Finally, we compared differential distributions obtained using the NWA with those computed using stable top quarks. To that end, we employed the same fixed scale choice in both calculations and reconstructed the top-quark kinematics from the decay products in the NWA. We find that large differences due to the fiducial phase-space cuts can be observed. At peak, the two computational approaches can reach differences of 60% in high p_T regimes, which are particularly relevant to BSM studies. Dimensionless quantities are also affected. For the rapidity of the

Higgs boson, we found that the calculation in the NWA is more centrally peaked. Notably, this effect is weakened for the $\mathcal{O}_{\phi G}$ and $\mathcal{O}_{tG}$ operators. We conclude that the SM behavior does not translate to the SMEFT case directly. Most of these shape differences are not covered by theoretical uncertainties and have to be carefully respected in experimental analyses. For some operators, differences to the SM of up to 20 percentage points can be found. We observed that these effects are largely constrained to the $\mathcal{O}_{\phi G}$ and $\mathcal{O}_{tG}$ operators, while $\mathcal{O}_{t\phi}$ is very similar to the SM in a large range of observables. Generally all deviations to the pure SM are strongly enhanced in quadratic terms, as one would expect.

Overall, we conclude that both NLO QCD corrections and realistic final state selections can have noticeable effects on SMEFT predictions and should be included in the corresponding studies where available. The regimes most sensitive to BSM physics often directly overlap with the regions in which we find the largest effects. At NLO QCD, missing higher order terms in the perturbative expansion contribute the largest share to the theoretical uncertainty of the calculation. Finally, the newly developed HELAC-SMEFT framework is well equipped to handle further analyses and provides a solid foundation for further theoretical advancements.

Bibliography

[1] ATLAS collaboration, *Measurement of the top-quark pole mass in dileptonic $t\bar{t}$ + 1-jet events at $\sqrt{s} = 13$ TeV with the ATLAS experiment*, *JHEP* **12** (2025) 023 [2507.02632].

[2] H.-Y. Bi, M. Kraus, M. Reinartz and M. Worek, *NLO QCD predictions for off-shell $t\bar{t}W$ production in association with a light jet at the LHC*, *JHEP* **09** (2023) 026 [2305.03802].

[3] G. Bevilacqua, M. Reinartz and M. Worek, *SMEFT everywhere: a NLO study of $\boldsymbol{pp \to t\bar{t}H}$ with decaying tops*, 2605.11747.

[4] M. Reinartz, *Off-shell $t\bar{t}W^+j$ production at NLO QCD accuracy*, *PoS* **ICHEP2024** (2025) 366 [2410.05959].

[5] ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1 [1207.7214].

[6] CMS collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30 [1207.7235].

[7] CMS collaboration, *A portrait of the Higgs boson by the CMS experiment ten years after the discovery.*, *Nature* **607** (2022) 60 [2207.00043].

[8] ATLAS collaboration, *A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery*, *Nature* **607** (2022) 52 [2207.00092].

[9] R. Balsach et al., *State-of-the-art cross sections for $t\bar{t}H$: NNLO predictions matched with NNLL resummation and EW corrections*, *SciPost* (2025) [2503.15043].

[10] CMS collaboration, *Evidence for the 125 GeV Higgs Boson Decaying to a Pair of τ Leptons*, *JHEP* **05** (2014) 104 [1401.5041].

[11] ATLAS collaboration, *Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector*, *JHEP* **04** (2015) 117 [1501.04943].

[12] ATLAS, 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*, *JHEP* **08** (2016) 045 [1606.02266].

[13] CMS collaboration, *Evidence for the Higgs boson decay to a bottom quark–antiquark pair*, *Phys. Lett. B* **780** (2018) 501 [1709.07497].

[14] ATLAS collaboration, *Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector*, *Phys. Lett. B* **786** (2018) 59 [1808.08238].

[15] CMS collaboration, *Observation of Higgs boson decay to bottom quarks*, *Phys. Rev. Lett.* **121** (2018) 121801 [1808.08242].

[16] CMS collaboration, *Evidence for associated production of a Higgs boson with a top quark pair in final states with electrons, muons, and hadronically decaying τ leptons at $\sqrt{s} = 13$ TeV*, *JHEP* **08** (2018) 066 [1803.05485].

[17] ATLAS collaboration, *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*, *Phys. Rev. D* **97** (2018) 072003 [1712.08891].

[18] ATLAS collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, *Phys. Lett. B* **784** (2018) 173 [1806.00425].

[19] CMS collaboration, *Observation of $\mathrm{t\bar{t}}H$ production*, *Phys. Rev. Lett.* **120** (2018) 231801 [1804.02610].

[20] CMS collaboration, *Evidence for Higgs boson decay to a pair of muons*, *JHEP* **01** (2021) 148 [2009.04363].

[21] CMS collaboration, *Search for Higgs Boson Decay to a Charm Quark-Antiquark Pair in Proton-Proton Collisions at s=13 TeV*, *Phys. Rev. Lett.* **131** (2023) 061801 [2205.05550].

[22] ATLAS collaboration, *Measurements of WH and ZH production with Higgs boson decays into bottom quarks and direct constraints on the charm Yukawa coupling in 13 TeV pp collisions with the ATLAS detector*, *JHEP* **04** (2025) 075 [2410.19611].

[23] Particle Data Group collaboration, *Review of particle physics*, *Phys. Rev. D* **110** (2024) 030001.

[24] ATLAS collaboration, *Observation of quantum entanglement with top quarks at the ATLAS detector*, *Nature* **633** (2024) 542 [2311.07288].

[25] J.H. Kühn, A. Scharf and P. Uwer, *Weak Interactions in Top-Quark Pair Production at Hadron Colliders: An Update*, *Phys. Rev. D* **91** (2015) 014020 [1305.5773].

[26] T. Martini and M. Schulze, *Electroweak loops as a probe of new physics in $t\bar{t}$ production at the LHC*, *JHEP* **04** (2020) 017 [1911.11244].

[27] T. Martini, R.-Q. Pan, M. Schulze and M. Xiao, *Probing the CP structure of the top quark Yukawa coupling: Loop sensitivity versus on-shell sensitivity*, *Phys. Rev. D* **104** (2021) 055045 [2104.04277].

[28] CMS collaboration, *Measurement of the top quark Yukawa coupling from $t\bar{t}$ kinematic distributions in the dilepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. D* **102** (2020) 092013 [2009.07123].

[29] Q.-H. Cao, S.-L. Chen and Y. Liu, *Probing Higgs Width and Top Quark Yukawa Coupling from $t\bar{t}H$ and $t\bar{t}t\bar{t}$ Productions*, *Phys. Rev. D* **95** (2017) 053004 [1602.01934].

[30] Q.-H. Cao, S.-L. Chen, Y. Liu, R. Zhang and Y. Zhang, *Limiting top quark-Higgs boson interaction and Higgs-boson width from multitop productions*, *Phys. Rev. D* **99** (2019) 113003 [1901.04567].

[31] CMS collaboration, *Search for production of four top quarks in final states with same-sign or multiple leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **80** (2020) 75 [1908.06463].

[32] K. Agashe et al., *Report of the Topical Group on Top quark physics and heavy flavor production for Snowmass 2021*, 2209.11267.

[33] F. Maltoni, D. Pagani and S. Tentori, *Top-quark pair production as a probe of light top-philic scalars and anomalous Higgs interactions*, *JHEP* **09** (2024) 098 [2406.06694].

[34] F. Demartin, F. Maltoni, K. Mawatari and M. Zaro, *Higgs production in association with a single top quark at the LHC*, *Eur. Phys. J. C* **75** (2015) 267 [1504.00611].

[35] LHC Higgs Cross Section Working Group collaboration, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, *CERN Yellow Rep. Monogr.* **2** (2017) 1 [1610.07922].

[36] C. collaboration, *Summary of run 2 tth signal strengths*, 2023.

[37] ATLAS collaboration, *CP Properties of Higgs Boson Interactions with Top Quarks in the $t\bar{t}H$ and tH Processes Using $H \to \gamma\gamma$ with the ATLAS Detector*, *Phys. Rev. Lett.* **125** (2020) 061802 [2004.04545].

[38] CMS collaboration, *Measurements of $t\bar{t}H$ Production and the CP Structure of the Yukawa Interaction between the Higgs Boson and Top Quark in the Diphoton Decay Channel*, *Phys. Rev. Lett.* **125** (2020) 061801 [2003.10866].

[39] CMS collaboration, *Search for CP violation in ttH and tH production in multilepton channels in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **07** (2023) 092 [2208.02686].

[40] ATLAS collaboration, *Measurement of the properties of Higgs boson production at $\sqrt{s} = 13$ TeV in the $H \to \gamma\gamma$ channel using 139 fb^{-1} of pp collision data with the ATLAS experiment*, *JHEP* **07** (2023) 088 [2207.00348].

[41] ATLAS collaboration, *Probing the CP nature of the top–Higgs Yukawa coupling in $tt^{-}H$ and tH events with $H \to bb^{-}$ decays using the ATLAS detector at the LHC*, *Phys. Lett. B* **849** (2024) 138469 [2303.05974].

[42] CMS collaboration, *Measurement of the $t\bar{t}H$ and tH production rates in the $H \to b\bar{b}$ decay channel using proton-proton collision data at $\sqrt{s} = 13$ TeV*, *JHEP* **02** (2025) 097 [2407.10896].

[43] W. Beenakker, S. Dittmaier, M. Kramer, B. Plumper, M. Spira and P.M. Zerwas, *Higgs radiation off top quarks at the Tevatron and the LHC*, *Phys. Rev. Lett.* **87** (2001) 201805 [hep-ph/0107081].

[44] W. Beenakker, S. Dittmaier, M. Kramer, B. Plumper, M. Spira and P.M. Zerwas, *NLO QCD corrections to t anti-t H production in hadron collisions*, *Nucl. Phys. B* **653** (2003) 151 [hep-ph/0211352].

[45] L. Reina, S. Dawson and D. Wackeroth, *QCD corrections to associated t anti-t h production at the Tevatron*, *Phys. Rev. D* **65** (2002) 053017 [hep-ph/0109066].

[46] L. Reina and S. Dawson, *Next-to-leading order results for t anti-t h production at the Tevatron*, *Phys. Rev. Lett.* **87** (2001) 201804 [hep-ph/0107101].

[47] S. Dawson, L.H. Orr, L. Reina and D. Wackeroth, *Associated top quark Higgs boson production at the LHC*, *Phys. Rev. D* **67** (2003) 071503 [hep-ph/0211438].

[48] S. Frixione, V. Hirschi, D. Pagani, H.S. Shao and M. Zaro, *Weak corrections to Higgs hadroproduction in association with a top-quark pair*, *JHEP* **09** (2014) 065 [1407.0823].

[49] S. Frixione, V. Hirschi, D. Pagani, H.S. Shao and M. Zaro, *Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons*, *JHEP* **06** (2015) 184 [1504.03446].

[50] Y. Zhang, W.-G. Ma, R.-Y. Zhang, C. Chen and L. Guo, *QCD NLO and EW NLO corrections to $t\bar{t}H$ production with top quark decays at hadron collider*, *Phys. Lett. B* **738** (2014) 1 [1407.1110].

[51] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H.S. Shao and M. Zaro, *The automation of next-to-leading order electroweak calculations*, *JHEP* **07** (2018) 185 [1804.10017].

[52] S. Catani, I. Fabre, M. Grazzini and S. Kallweit, *$t\bar{t}H$ production at NNLO: the flavour off-diagonal channels*, *Eur. Phys. J. C* **81** (2021) 491 [2102.03256].

[53] S. Catani, S. Devoto, M. Grazzini, S. Kallweit, J. Mazzitelli and C. Savoini, *Higgs Boson Production in Association with a Top-Antitop Quark Pair in Next-to-Next-to-Leading Order QCD*, *Phys. Rev. Lett.* **130** (2023) 111902 [2210.07846].

[54] S. Devoto, M. Grazzini, S. Kallweit, J. Mazzitelli and C. Savoini, *Precise predictions for $t\bar{t}H$ production at the LHC: inclusive cross section and differential distributions*, *JHEP* **03** (2025) 189 [2411.15340].

[55] F. Febres Cordero, G. Figueiredo, M. Kraus, B. Page and L. Reina, *Two-loop master integrals for leading-color $pp \to t\bar{t}H$ amplitudes with a light-quark loop*, *JHEP* **07** (2024) 084 [2312.08131].

[56] B. Agarwal, G. Heinrich, S.P. Jones, M. Kerner, S.Y. Klein, J. Lang et al., *Two-loop amplitudes for $t\bar{t}H$ production: the quark-initiated N_f-part*, *JHEP* **05** (2024) 013 [2402.03301].

[57] G. Wang, T. Xia, L.L. Yang and X. Ye, *Two-loop QCD amplitudes for $t\bar{t}H$ production from boosted limit*, *JHEP* **07** (2024) 121 [2402.00431].

[58] A. Kulesza, L. Motyka, T. Stebel and V. Theeuwes, *Soft gluon resummation for associated $t\bar{t}H$ production at the LHC*, *JHEP* **03** (2016) 065 [1509.02780].

[59] A. Broggio, A. Ferroglia, B.D. Pecjak, A. Signer and L.L. Yang, *Associated production of a top pair and a Higgs boson beyond NLO*, *JHEP* **03** (2016) 124 [1510.01914].

[60] A. Broggio, A. Ferroglia, B.D. Pecjak and L.L. Yang, *NNLL resummation for the associated production of a top pair and a Higgs boson at the LHC*, *JHEP* **02** (2017) 126 [1611.00049].

[61] A. Kulesza, L. Motyka, T. Stebel and V. Theeuwes, *Associated $t\bar{t}H$ production at the LHC: Theoretical predictions at NLO+NNLL accuracy*, *Phys. Rev. D* **97** (2018) 114007 [1704.03363].

[62] W.-L. Ju and L.L. Yang, *Resummation of soft and Coulomb corrections for $t\bar{t}h$ production at the LHC*, *JHEP* **06** (2019) 050 [1904.08744].

[63] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, R. Pittau and P. Torrielli, *Scalar and pseudoscalar Higgs production in association with a top–antitop pair*, *Phys. Lett. B* **701** (2011) 427 [1104.5613].

[64] H.B. Hartanto, B. Jager, L. Reina and D. Wackeroth, *Higgs boson production in association with top quarks in the POWHEG BOX*, *Phys. Rev. D* **91** (2015) 094003 [1501.04498].

[65] C. Biello, C. Savoini, C. Signorile-Signorile and M. Wiesemann, *Next-to-next-to-leading order event generation for $t\bar{t}H$ production with approximate two-loop amplitude*, 2603.06143.

[66] P.F. Monni, P. Nason, E. Re, M. Wiesemann and G. Zanderighi, *MiNNLO$_{PS}$: a new method to match NNLO QCD to parton showers*, *JHEP* **05** (2020) 143 [1908.06987].

[67] P.F. Monni, E. Re and M. Wiesemann, *MiNNLO$_{PS}$: optimizing $2 \to 1$ hadronic processes*, *Eur. Phys. J. C* **80** (2020) 1075 [2006.04133].

[68] T. Ježo and P. Nason, *On the Treatment of Resonances in Next-to-Leading Order Calculations Matched to a Parton Shower*, *JHEP* **12** (2015) 065 [1509.09071].

[69] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [hep-ph/0409146].

[70] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070 [0709.2092].

[71] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043 [1002.2581].

[72] A. Denner and R. Feger, *NLO QCD corrections to off-shell top-antitop production with leptonic decays in association with a Higgs boson at the LHC*, *JHEP* **11** (2015) 209 [1506.07448].

[73] A. Denner, J.-N. Lang, M. Pellen and S. Uccirati, *Higgs production in association with off-shell top-antitop pairs at NLO EW and QCD at the LHC*, *JHEP* **02** (2017) 053 [1612.07138].

[74] D. Stremmer and M. Worek, *Production and decay of the Higgs boson in association with top quarks*, *JHEP* **02** (2022) 196 [2111.01427].

[75] G. Degrassi, S. Di Vita, J. Elias-Miro, J.R. Espinosa, G.F. Giudice, G. Isidori et al., *Higgs mass and vacuum stability in the Standard Model at NNLO*, *JHEP* **08** (2012) 098 [1205.6497].

[76] F. Demartin, F. Maltoni, K. Mawatari, B. Page and M. Zaro, *Higgs characterisation at NLO in QCD: CP properties of the top-quark Yukawa interaction*, *Eur. Phys. J. C* **74** (2014) 3065 [1407.5089].

[77] J. Hermann, D. Stremmer and M. Worek, *$\mathcal{CP}$ structure of the top-quark Yukawa interaction: NLO QCD corrections and off-shell effects*, *JHEP* **09** (2022) 138 [2205.09983].

[78] J. Ellis, M. Madigan, K. Mimasu, V. Sanz and T. You, *Top, Higgs, Diboson and Electroweak Fit to the Standard Model Effective Field Theory*, *JHEP* **04** (2021) 279 [2012.02779].

[79] CMS collaboration, *Search for new physics in top quark production with additional leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV using effective field theory*, *JHEP* **03** (2021) 095 [2012.04120].

[80] CMS collaboration, *Probing effective field theory operators in the associated production of top quarks with a Z boson in multilepton final states at $\sqrt{s} = 13$ TeV*, *JHEP* **12** (2021) 083 [2107.13896].

[81] CMS collaboration, *Search for physics beyond the standard model in top quark production with additional leptons in the context of effective field theory*, *JHEP* **12** (2023) 068 [2307.15761].

[82] CMS collaboration, *Measurement of the inclusive and differential $t\bar{t}\gamma$ cross sections in the single-lepton channel and EFT interpretation at $\sqrt{s} = 13$ TeV*, *JHEP* **12** (2021) 180 [2107.01508].

[83] ATLAS collaboration, *Measurements of differential cross-sections in top-quark pair events with a high transverse momentum top quark and limits on beyond the Standard Model contributions to top-quark pair production with the ATLAS detector at $\sqrt{s}$ = 13 TeV*, *JHEP* **06** (2022) 063 [2202.12134].

[84] ATLAS collaboration, *Measurement of the polarisation of single top quarks and antiquarks produced in the t-channel at $\sqrt{s}$ = 13 TeV and bounds on the tWb dipole operator from the ATLAS experiment*, *JHEP* **11** (2022) 040 [2202.11382].

[85] ATLAS collaboration, *Inclusive and differential cross-section measurements of $t\bar{t}Z$ production in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector, including EFT and spin-correlation interpretations*, *JHEP* **07** (2024) 163 [2312.04450].

[86] ATLAS collaboration, *Measurements of inclusive and differential cross-sections of $t\bar{t}\gamma$ production in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, *JHEP* **10** (2024) 191 [2403.09452].

[87] CMS collaboration, *Search for CP violation in events with top quarks and Z bosons at $\sqrt{s}$ = 13 and 13.6 TeV*, *Phys. Lett. B* **869** (2025) 139857 [2505.21206].

[88] ATLAS collaboration, *Constraints on effective field theories via quadruple-differential angular decay rates from t-channel single-top-quark production at $\sqrt{s} = 13$ TeV with the ATLAS detector*, 2510.23372.

[89] ATLAS collaboration, *Top EFT summary plots April 2026*, .

[90] ATLAS collaboration, *Effective field theory interpretation of ATLAS measurements involving the Higgs boson, electroweak bosons and the top quark*, 2604.21670.

[91] C. Degrande, R. Rosenfeld and A. Vasquez, *Collider sensitivity to SMEFT heavy-quark operators at one-loop in top-quark processes*, *JHEP* **07** (2024) 114 [2402.06528].

[92] C. Zhang, *Single Top Production at Next-to-Leading Order in the Standard Model Effective Field Theory*, *Phys. Rev. Lett.* **116** (2016) 162002 [1601.06163].

[93] M. de Beurs, E. Laenen, M. Vreeswijk and E. Vryonidou, *Effective operators in t-channel single top production and decay*, *Eur. Phys. J. C* **78** (2018) 919 [1807.03576].

[94] T. Neumann and Z.E. Sullivan, *Off-Shell Single-Top-Quark Production in the Standard Model Effective Field Theory*, *JHEP* **06** (2019) 022 [1903.11023].

[95] 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*, *JHEP* **10** (2016) 123 [1607.05330].

[96] O. Bessidskaia Bylund, F. Maltoni, I. Tsinikos, E. Vryonidou and C. Zhang, *Probing top quark neutral couplings in the Standard Model Effective Field Theory at NLO in QCD*, *JHEP* **05** (2016) 052 [1601.08193].

[97] R. Aoude, H. El Faham, F. Maltoni and E. Vryonidou, *Complete SMEFT predictions for four top quark production at hadron colliders*, *JHEP* **10** (2022) 163 [2208.04962].

[98] I. Brivio, S. Bruggisser, F. Maltoni, R. Moutafis, T. Plehn, E. Vryonidou et al., *O new physics, where art thou? A global search in the top sector*, *JHEP* **02** (2020) 131 [1910.03606].

[99] N.P. Hartland, F. Maltoni, E.R. Nocera, J. Rojo, E. Slade, E. Vryonidou et al., *A Monte Carlo global analysis of the Standard Model Effective Field Theory: the top quark sector*, *JHEP* **04** (2019) 100 [1901.05965].

[100] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [1405.0301].

[101] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer and T. Reiter, *UFO - The Universal FeynRules Output*, *Comput. Phys. Commun.* **183** (2012) 1201 [1108.2040].

[102] L. Darmé et al., *UFO 2.0: the 'Universal Feynman Output' format*, *Eur. Phys. J. C* **83** (2023) 631 [2304.09883].

[103] I. Brivio, Y. Jiang and M. Trott, *The SMEFTsim package, theory and tools*, *JHEP* **12** (2017) 070 [1709.06492].

[104] I. Brivio, *SMEFTsim 3.0 — a practical guide*, *JHEP* **04** (2021) 073 [2012.11343].

[105] C. Degrande, G. Durieux, F. Maltoni, K. Mimasu, E. Vryonidou and C. Zhang, *Automated one-loop computations in the standard model effective field theory*, *Phys. Rev. D* **103** (2021) 096024 [2008.11743].

[106] J. Braun, B. Campillo Aveleira, G. Heinrich, M. Höfer, S.P. Jones, M. Kerner et al., *One-loop calculations in effective field theories with GoSam-3.0*, *SciPost Phys. Codeb.* **62** (2026) 1 [2507.23549].

[107] A. Dedes, W. Materkowska, M. Paraskevas, J. Rosiek and K. Suxho, *Feynman rules for the Standard Model Effective Field Theory in R_ξ -gauges*, *JHEP* **06** (2017) 143 [1704.03888].

[108] N.D. Christensen and C. Duhr, *FeynRules - Feynman rules made easy*, *Comput. Phys. Commun.* **180** (2009) 1614 [0806.4194].

[109] A. Alloul, N.D. Christensen, C. Degrande, C. Duhr and B. Fuks, *FeynRules 2.0 - A complete toolbox for tree-level phenomenology*, *Comput. Phys. Commun.* **185** (2014) 2250 [1310.1921].

[110] A. Dedes, J. Rosiek, M. Ryczkowski, K. Suxho and L. Trifyllis, *SmeftFR v3 – Feynman rules generator for the Standard Model Effective Field Theory*, *Comput. Phys. Commun.* **294** (2024) 108943 [2302.01353].

[111] E.E. Jenkins, A.V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators II: Yukawa Dependence*, *JHEP* **01** (2014) 035 [1310.4838].

[112] E.E. Jenkins, A.V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators I: Formalism and lambda Dependence*, *JHEP* **10** (2013) 087 [1308.2627].

[113] R. Alonso, E.E. Jenkins, A.V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators III: Gauge Coupling Dependence and Phenomenology*, *JHEP* **04** (2014) 159 [1312.2014].

[114] C. Duhr, A. Vasquez, G. Ventura and E. Vryonidou, *Two-loop renormalisation of quark and gluon fields in the SMEFT*, *JHEP* **07** (2025) 160 [2503.01954].

[115] S. Di Noi, B.A. Erdelyi and R. Gröber, *Complete two-loop Yukawa-induced running of the Higgs-gluon coupling in SMEFT*, 2510.14680.

[116] L. Born, J. Fuentes-Martín and A.E. Thomsen, *Next-to-Leading Order Running in the SMEFT*, 2601.19974.

[117] R. Aoude, F. Maltoni, O. Mattelaer, C. Severi and E. Vryonidou, *Renormalisation group effects on SMEFT interpretations of LHC data*, *JHEP* **09** (2023) 191 [2212.05067].

[118] S. Di Noi and R. Gröber, *Renormalisation group running effects in $pp \to t\bar{t}h$ in the Standard Model Effective Field Theory*, *Eur. Phys. J. C* **84** (2024) 403 [2312.11327].

[119] R. Bartocci, A. Biekötter and T. Hurth, *Renormalisation group evolution effects on global SMEFT analyses*, *JHEP* **05** (2025) 203 [2412.09674].

[120] S. Di Noi and R. Gröber, *Two loops, four tops and two $\gamma 5$ schemes: A renormalization story*, *Phys. Lett. B* **869** (2025) 139878 [2507.10295].

[121] S. Di Noi, H. El Faham, R. Gröber, M. Vitti and E. Vryonidou, *Constraining four-heavy-quark operators with top-quark, Higgs, and electroweak precision data*, *JHEP* **01** (2026) 025 [2507.01137].

[122] A. Celis, J. Fuentes-Martin, A. Vicente and J. Virto, *DsixTools: The Standard Model Effective Field Theory Toolkit*, *Eur. Phys. J. C* **77** (2017) 405 [1704.04504].

[123] J. Fuentes-Martin, P. Ruiz-Femenia, A. Vicente and J. Virto, *DsixTools 2.0: The Effective Field Theory Toolkit*, *Eur. Phys. J. C* **81** (2021) 167 [2010.16341].

[124] J. Aebischer, J. Kumar and D.M. Straub, *Wilson: a Python package for the running and matching of Wilson coefficients above and below the electroweak scale*, *Eur. Phys. J. C* **78** (2018) 1026 [1804.05033].

[125] S. Di Noi and L. Silvestrini, *RGESolver: a C++ library to perform renormalization group evolution in the Standard Model Effective Theory*, *Eur. Phys. J. C* **83** (2023) 200 [2210.06838].

[126] R. Boughezal, C.-Y. Chen, F. Petriello and D. Wiegand, *Top quark decay at next-to-leading order in the Standard Model Effective Field Theory*, *Phys. Rev. D* **100** (2019) 056023 [1907.00997].

[127] G. Bevilacqua, M. Czakon, M.V. Garzelli, A. van Hameren, A. Kardos, C.G. Papadopoulos et al., *HELAC-NLO*, *Comput. Phys. Commun.* **184** (2013) 986 [1110.1499].

[128] J. Campbell, J. Huston and F. Krauss, *The Black Book of Quantum Chromodynamics : a Primer for the LHC Era*, Oxford University Press (2018), 10.1093/oso/9780199652747.001.0001.

[129] D. Stremmer and M. Worek, *Complete NLO corrections to top-quark pair production with isolated photons*, *JHEP* **07** (2024) 091 [2403.03796].

[130] L. Mans, D. Stremmer and M. Worek, *Complete NLO corrections to full off-shell $t\bar{t}$ production in the $\ell + j$ decay channel*, 2512.18391.

[131] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys. B* **126** (1977) 298.

[132] V.N. Gribov and L.N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438.

[133] L.N. Lipatov, *The parton model and perturbation theory*, *Yad. Fiz.* **20** (1974) 181.

[134] Y.L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, *Sov. Phys. JETP* **46** (1977) 641.

[135] M. Constantinou, *The x-dependence of hadronic parton distributions: A review on the progress of lattice QCD*, *Eur. Phys. J. A* **57** (2021) 77 [2010.02445].

[136] HadStruc collaboration, *Accessing the Gluon Momentum Fraction of Nucleons through the Gradient Flow*, 2602.14260.

[137] NNPDF collaboration, *Parton distributions from high-precision collider data*, *Eur. Phys. J. C* **77** (2017) 663 [1706.00428].

[138] L.J. Dixon, *Calculating scattering amplitudes efficiently*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 95): QCD and Beyond*, pp. 539–584, 1, 1996 [hep-ph/9601359].

[139] L.J. Dixon, *A brief introduction to modern amplitude methods*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Particle Physics: The Higgs Boson and Beyond*, pp. 31–67, 2014, DOI [1310.5353].

[140] M. Cacciari, G.P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [0802.1189].

[141] S. Catani, Y.L. Dokshitzer, M.H. Seymour and B.R. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, *Nucl. Phys. B* **406** (1993) 187.

[142] Y.L. Dokshitzer, G.D. Leder, S. Moretti and B.R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001 [hep-ph/9707323].

[143] G. Bevilacqua, H.-Y. Bi, H.B. Hartanto, M. Kraus, M. Lupattelli and M. Worek, *$t\bar{t}b\bar{b}$ at the LHC: on the size of corrections and b-jet definitions*, *JHEP* **08** (2021) 008 [2105.08404].

[144] A. Banfi, G.P. Salam and G. Zanderighi, *Infrared safe definition of jet flavor*, *Eur. Phys. J. C* **47** (2006) 113 [hep-ph/0601139].

[145] A.J. Larkoski, *Flavor-changing nonglobal logarithms*, *Phys. Rev. D* **113** (2026) 034013 [2510.06031].

[146] T. Generet, *IRC-safe jet flavour without modifying anything*, 2511.23423.

[147] A. Behring et al., *Flavoured jet algorithms: a comparative study*, *JHEP* **09** (2025) 149 [2506.13449].

[148] D. Binosi and L. Theußl, *Jaxodraw: A graphical user interface for drawing feynman diagrams*, *Computer Physics Communications* **161** (2004) 76.

[149] A. Denner, S. Dittmaier, M. Roth and D. Wackeroth, *Predictions for all processes e+ e- —> 4 fermions + gamma*, *Nucl. Phys. B* **560** (1999) 33 [hep-ph/9904472].

[150] A. Denner, S. Dittmaier, M. Roth and L.H. Wieders, *Electroweak corrections to charged-current e+ e- —> 4 fermion processes: Technical details and further results*, *Nucl. Phys. B* **724** (2005) 247 [hep-ph/0505042].

[151] A. Denner and S. Dittmaier, *The Complex-mass scheme for perturbative calculations with unstable particles*, *Nucl. Phys. B Proc. Suppl.* **160** (2006) 22 [hep-ph/0605312].

[152] A. Denner and J.-N. Lang, *The Complex-Mass Scheme and Unitarity in perturbative Quantum Field Theory*, *Eur. Phys. J. C* **75** (2015) 377 [1406.6280].

[153] M.D. Schwartz, *Quantum Field Theory and the Standard Model*, Cambridge University Press (3, 2014), 10.1017/9781139540940.

[154] I. Brivio and M. Trott, *The Standard Model as an Effective Field Theory*, *Phys. Rept.* **793** (2019) 1 [1706.08945].

[155] M. Beneke and V.A. Smirnov, *Asymptotic expansion of Feynman integrals near threshold*, *Nucl. Phys. B* **522** (1998) 321 [hep-ph/9711391].

[156] V.A. Smirnov, *Applied asymptotic expansions in momenta and masses*, *Springer Tracts Mod. Phys.* **177** (2002) 1.

[157] B. Jantzen, *Foundation and generalization of the expansion by regions*, *JHEP* **12** (2011) 076 [1111.2589].

[158] A.V. Manohar, *Introduction to Effective Field Theories*, 1804.05863.

[159] M.K. Gaillard, *The Effective One Loop Lagrangian With Derivative Couplings*, *Nucl. Phys. B* **268** (1986) 669.

[160] O. Cheyette, *Effective Action for the Standard Model With Large Higgs Mass*, *Nucl. Phys. B* **297** (1988) 183.

[161] B. Henning, X. Lu and H. Murayama, *How to use the Standard Model effective field theory*, *JHEP* **01** (2016) 023 [1412.1837].

[162] A. Drozd, J. Ellis, J. Quevillon and T. You, *The Universal One-Loop Effective Action*, *JHEP* **03** (2016) 180 [1512.03003].

[163] B. Henning, X. Lu and H. Murayama, *One-loop Matching and Running with Covariant Derivative Expansion*, *JHEP* **01** (2018) 123 [1604.01019].

[164] S.A.R. Ellis, J. Quevillon, T. You and Z. Zhang, *Mixed heavy–light matching in the Universal One-Loop Effective Action*, *Phys. Lett. B* **762** (2016) 166 [1604.02445].

[165] S.A.R. Ellis, J. Quevillon, T. You and Z. Zhang, *Extending the Universal One-Loop Effective Action: Heavy-Light Coefficients*, *JHEP* **08** (2017) 054 [1706.07765].

[166] Z. Zhang, *Covariant diagrams for one-loop matching*, *JHEP* **05** (2017) 152 [1610.00710].

[167] R. Pittau, *Matching high-energy electroweak fermion loops onto the Fermi theory without higher dimensional operators*, *Nucl. Phys. B* **950** (2020) 114835 [1902.01767].

[168] A.J. Buras and P.H. Weisz, *QCD Nonleading Corrections to Weak Decays in Dimensional Regularization and 't Hooft-Veltman Schemes*, *Nucl. Phys. B* **333** (1990) 66.

[169] S. Herrlich and U. Nierste, *Evanescent operators, scheme dependences and double insertions*, *Nucl. Phys. B* **455** (1995) 39 [hep-ph/9412375].

[170] J. Fuentes-Martín, M. König, J. Pagès, A.E. Thomsen and F. Wilsch, *Evanescent operators in one-loop matching computations*, *JHEP* **02** (2023) 031 [2211.09144].

[171] M.J. Dugan and B. Grinstein, *On the vanishing of evanescent operators*, *Phys. Lett. B* **256** (1991) 239.

[172] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085 [1008.4884].

[173] W. Buchmuller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, *Nucl. Phys. B* **268** (1986) 621.

[174] H.-L. Li, Z. Ren, M.-L. Xiao, J.-H. Yu and Y.-H. Zheng, *Operators for generic effective field theory at any dimension: on-shell amplitude basis construction*, *JHEP* **04** (2022) 140 [2201.04639].

[175] R.V. Harlander and M.C. Schaaf, *AutoEFT: Automated operator construction for effective field theories*, *Comput. Phys. Commun.* **300** (2024) 109198 [2309.15783].

[176] S. Weinberg, *Baryon and Lepton Nonconserving Processes*, *Phys. Rev. Lett.* **43** (1979) 1566.

[177] A. Giarnetti, J. Herrero-Garcia, S. Marciano, D. Meloni and D. Vatsyayan, *Neutrino masses from new Weinberg-like operators: phenomenology of TeV scalar multiplets*, *JHEP* **05** (2024) 055 [2312.13356].

[178] P. Minkowski, $\mu \to e\gamma$ *at a rate of one out of 109 muon decays?*, *Physics Letters B* **67** (1977) 421.

[179] T. Yanagida, *Horizontal Symmetry and Masses of Neutrinos*, *Prog. Theor. Phys.* **64** (1980) 1103.

[180] Y. Liao and X.-D. Ma, *Renormalization Group Evolution of Dimension-seven Baryon- and Lepton-number-violating Operators*, *JHEP* **11** (2016) 043 [1607.07309].

[181] L. Lehman, *Extending the Standard Model Effective Field Theory with the Complete Set of Dimension-7 Operators*, *Phys. Rev. D* **90** (2014) 125023 [1410.4193].

[182] H.-L. Li, Z. Ren, J. Shu, M.-L. Xiao, J.-H. Yu and Y.-H. Zheng, *Complete set of dimension-eight operators in the standard model effective field theory*, *Phys. Rev. D* **104** (2021) 015026 [2005.00008].

[183] H.-L. Li, Z. Ren, M.-L. Xiao, J.-H. Yu and Y.-H. Zheng, *Complete set of dimension-nine operators in the standard model effective field theory*, *Phys. Rev. D* **104** (2021) 015025 [2007.07899].

[184] R.V. Harlander, T. Kempkens and M.C. Schaaf, *Standard model effective field theory up to mass dimension 12*, *Phys. Rev. D* **108** (2023) 055020 [2305.06832].

[185] C. Degrande and J. Touchèque, *A reduced basis for CP violation in SMEFT at colliders and its application to diboson production*, *JHEP* **04** (2022) 032 [2110.02993].

[186] G. Passarino, *XEFT, the challenging path up the hill: dim = 6 and dim = 8*, 1901.04177.

[187] J.M. Cullen, B.D. Pecjak and D.J. Scott, *NLO corrections to $h \to b\bar{b}$ decay in SMEFT*, *JHEP* **08** (2019) 173 [1904.06358].

[188] A. Denner and S. Dittmaier, *Electroweak Radiative Corrections for Collider Physics*, *Phys. Rept.* **864** (2020) 1 [1912.06823].

[189] M. Misiak, M. Paraskevas, J. Rosiek, K. Suxho and B. Zglinicki, *Effective Field Theories in R_ξ gauges*, *JHEP* **02** (2019) 051 [1812.11513].

[190] A. Helset, M. Paraskevas and M. Trott, *Gauge fixing the Standard Model Effective Field Theory*, *Phys. Rev. Lett.* **120** (2018) 251801 [1803.08001].

[191] A. Helset, A. Martin and M. Trott, *The Geometric Standard Model Effective Field Theory*, *JHEP* **03** (2020) 163 [2001.01453].

[192] D. Barducci et al., *Interpreting top-quark LHC measurements in the standard-model effective field theory*, 1802.07237.

[193] F. Krauss, S. Kuttimalai and T. Plehn, *LHC multijet events as a probe for anomalous dimension-six gluon interactions*, *Phys. Rev. D* **95** (2017) 035024 [1611.00767].

[194] A. van Hameren, C.G. Papadopoulos and R. Pittau, *Automated one-loop calculations: A Proof of concept*, *JHEP* **09** (2009) 106 [0903.4665].

[195] M. Czakon, C.G. Papadopoulos and M. Worek, *Polarizing the Dipoles*, *JHEP* **08** (2009) 085 [0905.0883].

[196] G. Bevilacqua, M. Czakon, M. Kubocz and M. Worek, *Complete Nagy-Soper subtraction for next-to-leading order calculations in QCD*, *JHEP* **10** (2013) 204 [1308.5605].

[197] A. van Hameren, *Kaleu: A General-Purpose Parton-Level Phase Space Generator*, 1003.4953.

[198] A. van Hameren, *PARNI for importance sampling and density estimation*, *Acta Phys. Polon. B* **40** (2009) 259 [0710.2448].

[199] A. Kanaki and C.G. Papadopoulos, *HELAC: A Package to compute electroweak helicity amplitudes*, *Comput. Phys. Commun.* **132** (2000) 306 [hep-ph/0002082].

[200] G. Bevilacqua, "HEPlot.".

[201] G. Ossola, C.G. Papadopoulos and R. Pittau, *Reducing full one-loop amplitudes to scalar integrals at the integrand level*, *Nucl. Phys. B* **763** (2007) 147 [hep-ph/0609007].

[202] G. Ossola, C.G. Papadopoulos and R. Pittau, *CutTools: A Program implementing the OPP reduction method to compute one-loop amplitudes*, *JHEP* **03** (2008) 042 [0711.3596].

[203] A. van Hameren, *OneLOop: For the evaluation of one-loop scalar functions*, *Comput. Phys. Commun.* **182** (2011) 2427 [1007.4716].

[204] C.G. Papadopoulos and M. Worek, *Multi-parton cross sections at hadron colliders*, *Eur. Phys. J. C* **50** (2007) 843 [hep-ph/0512150].

[205] P.D. Draggiotis, R.H.P. Kleiss and C.G. Papadopoulos, *Multijet production in hadron collisions*, *Eur. Phys. J. C* **24** (2002) 447 [hep-ph/0202201].

[206] A. Cafarella, C.G. Papadopoulos and M. Worek, *Helac-Phegas: A Generator for all parton level processes*, *Comput. Phys. Commun.* **180** (2009) 1941 [0710.2427].

[207] F. Maltoni, K. Paul, T. Stelzer and S. Willenbrock, *Color Flow Decomposition of QCD Amplitudes*, *Phys. Rev. D* **67** (2003) 014026 [hep-ph/0209271].

[208] M.L. Mangano, *Introduction to QCD*, in *1998 European School of High-Energy Physics*, pp. 53–97, 1998.

[209] D. Stremmer, *Precise modeling of prompt photons in associated top-quark pair production*, Ph.D. thesis, RWTH Aachen University, RWTH Aachen U., 2024. 10.18154/RWTH-2024-08188.

[210] G. Ossola, C.G. Papadopoulos and R. Pittau, *Numerical evaluation of six-photon amplitudes*, *JHEP* **07** (2007) 085 [0704.1271].

[211] W.L. van Neerven and J.A.M. Vermaseren, *LARGE LOOP INTEGRALS*, *Phys. Lett. B* **137** (1984) 241.

[212] E. Remiddi and L. Tancredi, *Schouten identities for Feynman graph amplitudes; The Master Integrals for the two-loop massive sunrise graph*, *Nucl. Phys. B* **880** (2014) 343 [1311.3342].

[213] T. Binoth, J.P. Guillet and G. Heinrich, *Reduction formalism for dimensionally regulated one loop N point integrals*, *Nucl. Phys. B* **572** (2000) 361 [hep-ph/9911342].

[214] R.K. Ellis, Z. Kunszt, K. Melnikov and G. Zanderighi, *One-loop calculations in quantum field theory: from Feynman diagrams to unitarity cuts*, *Phys. Rept.* **518** (2012) 141 [1105.4319].

[215] G. Ossola, C.G. Papadopoulos and R. Pittau, *On the Rational Terms of the one-loop amplitudes*, *JHEP* **05** (2008) 004 [0802.1876].

[216] A. Lazopoulos, K. Melnikov and F. Petriello, *QCD corrections to tri-boson production*, *Phys. Rev. D* **76** (2007) 014001 [hep-ph/0703273].

[217] T. Binoth, G. Ossola, C.G. Papadopoulos and R. Pittau, *NLO QCD corrections to tri-boson production*, *JHEP* **06** (2008) 082 [0804.0350].

[218] J. Alwall et al., *A Standard format for Les Houches event files*, *Comput. Phys. Commun.* **176** (2007) 300 [hep-ph/0609017].

[219] S. Catani and M.H. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, *Nucl. Phys. B* **485** (1997) 291 [hep-ph/9605323].

[220] S. Catani, S. Dittmaier, M.H. Seymour and Z. Trocsanyi, *The Dipole formalism for next-to-leading order QCD calculations with massive partons*, *Nucl. Phys. B* **627** (2002) 189 [hep-ph/0201036].

[221] Z. Nagy, *Next-to-leading order calculation of three jet observables in hadron hadron collision*, *Phys. Rev. D* **68** (2003) 094002 [hep-ph/0307268].

[222] J.M. Campbell, R.K. Ellis and F. Tramontano, *Single top production and decay at next-to-leading order*, *Phys. Rev. D* **70** (2004) 094012 [hep-ph/0408158].

[223] PDF4LHC Working Group collaboration, *The PDF4LHC21 combination of global PDF fits for the LHC Run III*, *J. Phys. G* **49** (2022) 080501 [2203.05506].

[224] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht et al., *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J. C* **75** (2015) 132 [1412.7420].

[225] J.A.M. Vermaseren, *New features of FORM*, math-ph/0010025.

[226] I. Antcheva et al., *ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization*, *Comput. Phys. Commun.* **180** (2009) 2499 [1508.07749].

[227] Z. Bern, L.J. Dixon, F. Febres Cordero, S. Höche, H. Ita, D.A. Kosower et al., *Ntuples for NLO Events at Hadron Colliders*, *Comput. Phys. Commun.* **185** (2014) 1443 [1310.7439].

[228] S. Badger, B. Biedermann and P. Uwer, *NGluon: A Package to Calculate One-loop Multi-gluon Amplitudes*, *Comput. Phys. Commun.* **182** (2011) 1674 [1011.2900].

[229] J. Hermann, *Signal and background modeling in BSM studies in the top-quark sector at the LHC*, Ph.D. thesis, RWTH Aachen University, RWTH Aachen U., 2023.

[230] M. Jezabek and J.H. Kuhn, *QCD Corrections to Semileptonic Decays of Heavy Quarks*, *Nucl. Phys. B* **314** (1989) 1.