TECHNISCHE UNIVERSITÄT MÜNCHEN
TUM SCHOOL OF NATURAL SCIENCES

# PRECISION LHC PHENOMENOLOGY VIA NNLO PREDICTIONS MATCHED TO PARTON SHOWERS

## SILVIA ZANOLI

Vollständiger Abdruck der von der TUM School of Natural Sciences der Technischen Universität München zur Erlangung des akademischen Grades einer

**Doktorin der Naturwissenschaften (Dr. rer. nat.)**

genehmigten Dissertation.

| | |
|---|---|
| Vorsitz: | Prof. Dr. Laura Fabbietti |
| Prüfer*innen der Dissertation: | 1. Prof. Dr. Andreas Weiler |
| | 2. Prof. Dr. Giulia Zanderighi |

Die Dissertation wurde am 21.06.2023 bei der Technischen Universität München eingereicht und durch die TUM School of Natural Sciences am 07.09.2023 angenommen.

MAX-PLANCK-INSTITUT FÜR PHYSIK

# PRECISION LHC PHENOMENOLOGY VIA NNLO PREDICTIONS MATCHED TO PARTON SHOWERS

**SILVIA ZANOLI**

**Novel Computational Techniques in Particle Physics**
Director: Prof. Dr. Giulia Zanderighi

The doctoral studies were supported by the International Max Planck Research School (IMPRS) Fellowship on Elementary Particle Physics.

*Ai miei nonni,*
*che ancora mi chiedono*
*che cosa io stia ricercando.*

*A mamma e papá,*
*che anche se non me lo chiedono*
*lo pensano lo stesso.*

# Abstract

In light of the increasing level of precision of experimental measurements, providing theoretical predictions at the highest accuracy possible is crucial for a correct understanding of current and future LHC data. In this thesis we analyse one of the main building blocks to obtain realistic and precise simulations of events at colliders, namely the matching between fixed-order calculations and parton showers.

Fixed-order calculations in perturbation theory and parton showers are the two main tools needed to describe an event at a collider. The former is suitable to describe accurately the hard scattering process, the latter provides a realistic picture in the soft and collinear approximation of the production of a high-multiplicity final state. They are thus two complementary approaches that need to be combined in order to achieve the most realistic predictions. The current state of the art for such matching procedure is given by NNLO+PS accuracy, which means retaining the NNLO QCD accuracy of the fixed-order computation without spoiling the structure of the parton shower.

The MiNNLO$_{\text{PS}}$ method is a powerful tool for reaching NNLO+PS accuracy, as it requires neither a computationally intense reweighting nor the introduction of unphysical parameters for partitioning the phase space. In this thesis we present cutting-edge results obtained by applying this method to relevant LHC processes.

We first present a phenomenological analysis for Higgsstrahlung with subsequent decay of the Higgs boson to a pair of bottom quarks, which is crucial for the determination of the bottom Yukawa coupling. We show that the inclusion of NNLO QCD effects is fundamental for a correct description of this process. Moreover, we present non-trivial effects associated to the identification of $b$-jets through different jet-clustering algorithms.

We then extend the computation of $pp \to ZH \to \ell^+\ell^- b\bar{b}$ production to the SM Effective Field Theory (SMEFT), including dimension-six operators in the calculation that modify the SM Higgs field and couplings up to NNLO QCD accuracy, targeting NNLO+PS accuracy in the SMEFT. We present kinematic observables that are strongly affected by the considered operators and can thus be used to enhance our sensitivity to new physics effects.

We conclude by presenting a novel strategy to match NNLO QCD and NLO EW computations with parton showers using an a-posteriori reweighting. We discuss phenomenological results for $W^\pm Z$ production, which can be used as a probe of the gauge symmetry structure of the EW sector. We identify different combination schemes between QCD and EW computations matched with QCD and/or QED showers and we analyse their capabilities to describe certain kinematic regions of relevant distributions.

# Zusammenfassung

Angesichts der zunehmenden Präzision experimenteller Messungen ist es notwendig, theoretische Vorhersagen mit größtmöglicher Genauigkeit zu erreichen, um ein akkurates Verständnis der aktuellen und zukünftigen LHC-Daten zu gewährleisten. In dieser Arbeit beleuchten wir einen der wesentlichen Bausteine um eine realistische und präzise Simulation von Streuereignissen an Teilchenbeschleunigern zu erreichen, nämlich die Kombination von Präzisionsrechnungen zu fester Ordnung in der Störungstheorie und Teilchenschauern.

Die Berechnung höherer Ordnungen in der Störungstheorie sowie Teilchenschauer sind die beiden Hauptbestandteile, die benötigt werden, um Ereignisse an Teilchenbeschleunigern zu beschreiben. Ersteres dient der genauen Beschreibung des harten Streuprozesses, während letzteres eine realistische Beschreibung für die Produktion von Endzuständen mit hoher Teilchenzahl im Bereich der niederenergetischen und kollinearen Näherung liefert. Es handelt sich somit um zwei komplementäre Ansätze, die zusammengeführt werden müssen, um ihr volles Potenzial auszuschöpfen. Der aktuelle Stand der Wissenschaft für solch eine Kombination entspricht der Genauigkeit von NNLO+PS, d.h. der Beibehaltung der NNLO-Genauigkeit in Störungstheorie ohne Beeinträchtigung der Struktur des Teilchenschauers.

Die MINNLOPS-Methode ist ein wichtiges Werkzeug, um die NNLO+PS-Genauigkeit zu erreichen, da sie weder eine rechenintensive Neugewichtung der simulierten Streuereignisse noch die Einführung unphysikalischer Parameter zur Unterteilung des Phasenraums erfordert. In dieser Arbeit präsentieren wir bahnbrechende Ergebnisse, die mit Hilfe dieser Methode für wichtige LHC-Prozesse erzielt wurden.

Zunächst präsentieren wir eine phänomenologische Analyse für die Higgsstrahlung mit anschließendem Zerfall des Higgsteilchens in ein Bottom-Quark-Paar, was entscheidend für die Bestimmung der Yukawa-Kopplung von Bottom-Quarks ist. Wir zeigen, dass die Einbeziehung von NNLO-Effekten für eine korrekte Beschreibung dieses Prozesses unerlässlich ist. Darüber hinaus präsentieren wir nicht-triviale Effekte, die mit der Identifizierung von $b$-Jets durch verschiedene Clusterbildungsalgorithmen von Jets verbunden sind.

Darauf folgend erweitern wir die Berechnung des Streuprozesses $pp \to ZH \to \ell^+\ell^- b\bar{b}$ auf die Standardmodell-Effektive-Feldtheorie (SMEFT), indem wir die Beiträge von Operatoren bis Dimension sechs in die Rechnung aufnehmen, die das Higgsfeld des Standardmodells und die Kopplungen bis zur Ordnung NNLO in der Quantenchromodynamik (QCD) modifizieren, mit dem Ziel eine NNLO+PS Rechnung in der SMEFT zur erreichen. Wir präsentieren kinematische Observablen, die stark von den betrachteten Operatoren beeinflusst werden und daher unsere Sensitivität erhöhen können um Effekte durch neue Physik zu finden.

Abschließend präsentieren wir eine neuartige Strategie um NNLO-QCD- und NLO-EW-Berechnungen mit Teilchenschauern durch eine nachträgliche Neugewichtung der jeweiligen Streuereignisse zu kombinieren. Wir diskutieren phänomenologische Ergebnisse für die

$W^{\pm}Z$-Produktion, welche als Test der Eichstruktur des EW-Sektors dienen können. Wir identifizieren verschiedene Kombinationsschemata der QCD- und EW-Korrekturen, welche konsistent mit QCD- und/oder QED-Schauern kombiniert sind, und analysieren ihre Fähigkeit bestimmte kinematische Bereiche relevanter Verteilungen zu beschreiben.

# Contents

**1. Introduction** **2**

## PART I: THE THEORY

**2. Theoretical Framework** **10**
- 2.1. Fixed-order calculations . . . . . 10
  - 2.1.1. Preliminaries . . . . . 10
  - 2.1.2. Ultraviolet divergences . . . . . 13
  - 2.1.3. Infrared divergences . . . . . 17
  - 2.1.4. Subtraction method . . . . . 25
- 2.2. Parton Showers . . . . . 29

**3. Matching fixed-order calculations with parton showers** **34**
- 3.1. The POWHEG method . . . . . 35
- 3.2. The MiNLO′ method . . . . . 38
- 3.3. The MiNNLO$_{\text{PS}}$ method . . . . . 42

## PART II: PRECISION LHC PHENOMENOLOGY

**4. Higgsstrahlung with $H \to b\bar{b}$ decay at NNLO+PS accuracy in the SM** **54**
- 4.1. Motivation . . . . . 54
- 4.2. Outline of the calculation . . . . . 56
  - 4.2.1. Generation of production events . . . . . 57
  - 4.2.2. Generation of decay events . . . . . 58
  - 4.2.3. Combination of production and decay events . . . . . 59
- 4.3. Validation of the calculation . . . . . 60
  - 4.3.1. $W^{\pm}H$ production . . . . . 60
  - 4.3.2. $ZH$ production with $H \to b\bar{b}$ decay . . . . . 62
- 4.4. Phenomenological results . . . . . 65
  - 4.4.1. Input and parameters . . . . . 65
  - 4.4.2. Inclusive and fiducial cross sections . . . . . 66
  - 4.4.3. Differential distributions . . . . . 67
  - 4.4.4. Comparison to data . . . . . 69
- 4.5. Jet flavour . . . . . 73
  - 4.5.1. Infrared-safe definition of jet flavour . . . . . 73

4.5.2. Impact of the jet-clustering algorithm . . . . . . . . . . . . . . . . . . . . 75

**5. $ZH$ production with $H \to b\bar{b}$ decay at NNLO+PS accuracy in the SMEFT 80**
5.1. Motivation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 80
5.2. Outline of the calculation . . . . . . . . . . . . . . . . . . . . . . . . . . . 81
5.2.1. SMEFT factorizable contributions to $H \to b\bar{b}$ decay . . . . . . . . . . . 83
5.2.2. SMEFT non-factorizable contributions to $H \to b\bar{b}$ decay . . . . . . . . . 84
5.2.3. SMEFT contributions to $pp \to ZH$ . . . . . . . . . . . . . . . . . . . . 85
5.3. NNLO+PS implementation in the SMEFT . . . . . . . . . . . . . . . . . . . 86
5.3.1. Generation of events . . . . . . . . . . . . . . . . . . . . . . . . . . . 86
5.3.2. Combination of production and decay events . . . . . . . . . . . . . . . 87
5.3.3. NNLO+PS SMEFT differential cross section . . . . . . . . . . . . . . . . 88
5.4. Phenomenological results . . . . . . . . . . . . . . . . . . . . . . . . . . . 89
5.4.1. Input and parameters . . . . . . . . . . . . . . . . . . . . . . . . . . 89
5.4.2. Inclusive $H \to b\bar{b}$ decay . . . . . . . . . . . . . . . . . . . . . . . . . 90
5.4.3. Differential distributions . . . . . . . . . . . . . . . . . . . . . . . . 90

**6. $WZ$ production at NNLO QCD and NLO EW accuracy matched to parton showers 94**
6.1. Motivation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 94
6.2. Outline of the calculation . . . . . . . . . . . . . . . . . . . . . . . . . . . 97
6.2.1. Generation of events . . . . . . . . . . . . . . . . . . . . . . . . . . . 98
6.2.2. Matching with parton showers . . . . . . . . . . . . . . . . . . . . . . 98
6.2.3. Combination of QCD and EW results . . . . . . . . . . . . . . . . . . . 99
6.3. Validation of the calculation . . . . . . . . . . . . . . . . . . . . . . . . . 101
6.3.1. Input and parameters . . . . . . . . . . . . . . . . . . . . . . . . . . 101
6.3.2. NNLO QCD . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 102
6.3.3. NLO EW . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 102
6.4. Phenomenological results . . . . . . . . . . . . . . . . . . . . . . . . . . . 102
6.4.1. Settings . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 102
6.4.2. Differential distributions . . . . . . . . . . . . . . . . . . . . . . . . 105
6.4.3. Comparison to data . . . . . . . . . . . . . . . . . . . . . . . . . . . 110

## PART III: CONCLUSIONS

**7. Conclusions 116**

## APPENDIX

**A. Differential distributions for $W^-H$ production with $H \to b\bar{b}$ decay at NNLO+PS in the SM 122**

**Bibliography 128**

# List of Figures

2.1. Feynman diagrams contributing up to NLO accuracy in the process $e^+e^- \to q\bar{q}$. Panel (a) shows the LO diagram, panels (b) and (c) are the real corrections and (d) is the virtual contribution. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17
2.2. Examples of Feynman diagrams for proton-proton collisions where a real (left) and virtual (right) gluon is emitted from the initial state. The cancellation between the two contributions holds only in the soft limit ($z \to 1$), as the momentum of the parton undergoing the hard scattering is unaltered. When we consider the collinear limit, the parton momentum changes ($\vec{p} \to z\vec{p}$) in the real contribution, while it remains the same in the virtual one ($\vec{p}$). This mismatch leads to a non-complete cancellation of collinear singularities. . . . . 23

4.1. Main production channels of the Higgs boson at the LHC. Panel (a) is gluon-gluon fusion; panel (b) is vector-boson fusion; panel (c) is Higgsstrahlung; panel (d) is associated production with a pair of top quarks. . . . . . . . . . . . 55
4.2. Feynman diagrams for $VH$ production with subsequent $H \to b\bar{b}$ decay. Panels (a) and (b) are $q\bar{q}$-induced (Drell-Yan like) channels at LO, while panel (c) is a $gg$-induced channel that enters the calculation at NNLO. . . . . . . . . . . . . . 56
4.3. Differential distributions for $e^+\nu_e H$ production obtained with the MiNNLO$_{\mathrm{PS}}$ and the NNLOPS methods at LHE level. . . . . . . . . . . . . . . . . . . . . . . . 61
4.4. Differential ditributions for $e^+e^-b\bar{b}$ production obtained with the MiNNLO$_{\mathrm{PS}}$ and the NNLOPS methods at LHE level in the fiducial setup reported in table 4.3. 64
4.5. Invariant mass $m_{b\bar{b}}$ of the $b\bar{b}$ system for $e^+\nu_e b\bar{b}$ production (left) and for $e^+e^-b\bar{b}$ production (right) in the `fiducial-YR` setup. . . . . . . . . . . . . . . . . . . . . 68
4.6. Absolute difference in rapidity $|\Delta y_{b,b}|$ of the $b\bar{b}$ system for $e^+\nu_e b\bar{b}$ production (left) and for $e^+e^-b\bar{b}$ production (right) in the `fiducial-YR` setup. . . . . . . . . 69
4.7. Transverse momentum of the $b\bar{b}$ system $p_{\mathrm{T},bb}$ and of the leading $b$-jet $p_{\mathrm{T},b_{\mathrm{hard}}}$ for $e^+\nu_e b\bar{b}$ (left) and $e^+e^-b\bar{b}$ production (right) in the `fiducial-YR` setup. . . . . . 70
4.8. Transverse momentum of the electron/positron $p_{\mathrm{T},e^+}/p_{\mathrm{T},e^-}$ and invariant mass of the colour singlet $m_{WH}/m_{ZH}$ for $e^+\nu_e b\bar{b}$ (left) and for $e^+e^-b\bar{b}$ production (right) in the `fiducial-YR` setup. . . . . . . . . . . . . . . . . . . . . . . . . . . 71
4.9. Comparison of MiNNLO$_{\mathrm{PS}}$ predictions to ATLAS data in the `fiducial-ATLAS` setup. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 72
4.10. Possible $\mathcal{O}(\alpha_s^2)$ configuration in which a pair of widely separated soft quarks ($k_3$ and $k_4$) coming from a large-angle soft gluon spoils the flavour of jets ($k_1$ and $k_2$). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 73

4.11. Invariant mass $m_{bb}$ and transverse momentum $p_{\mathrm{T},bb}$ of the $b\bar{b}$ system for $e^+\nu_e b\bar{b}$ production in the `fiducial-YR` setup using the anti-$k_T$, $k_T$ and flavour-$k_T$ algorithms with $R = 0.4$. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 76
4.12. Transverse momentum of the leading $b$-jet $p_{\mathrm{T},b_{\mathrm{hard}}}$ and of the positron $p_{\mathrm{T},e^+}$ for $e^+\nu_e b\bar{b}$ production in the `fiducial-YR` setup using the anti-$k_T$, $k_T$ and flavour-$k_T$ algorithms with $R = 0.4$. . . . . . . . . . . . . . . . . . . . . . . . . . 77

5.1. Sample Feynman diagrams contributing to the $H \to b\bar{b}$ decay up to N$^3$LO accuracy with the insertion of the operator $Q_{bG}$. . . . . . . . . . . . . . . . . . . . 85
5.2. Sample of interfering diagrams associated to SMEFT corrections proportional to $c_{HG}$. They contribute to the partial decay width $H \to gg$ needed in (5.20). The left diagram is the LO SMEFT result with the insertion of the $Q_{HG}$ operator, while the right diagram is the one-loop SM contribution mediated by a top-quark loop. In the latter case, there are also contributions mediated by light quarks but they have a negligible numerical impact. . . . . . . . . . . . . . . . . 89
5.3. Invariant mass of the two reconstructed $b$-jets in the fiducial setup reported in table 5.1 for $c_{bH} = 0.15$ (left panel) and $c_{bG} = 400$ (right panel). The SMEFT prediction is represented by the red curve. The SM result with its scale uncertainty band is shown in black and gray. The radius used in the clustering algorithm (anti-$k_T$) is $R = 0.4$. . . . . . . . . . . . . . . . . . . . . . . . . . . 92
5.4. Invariant mass of the two reconstructed $b$-jets in the fiducial setup reported in table 5.1 for $c_{bG} = 400$. The SMEFT prediction is represented by the red curve. The SM result with its scale uncertainty band is shown in black and gray. The radii used in the clustering algorithm (anti-$k_T$) are $R = 0.7$ (left panel) and $R = 1$ (right panel). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 93
5.5. Invariant mass of the $b\bar{b}j$ system in the fiducial setup reported in table 5.1 for $c_{bG} = 400$. The SMEFT prediction is represented by the red curve. The SM result with its scale uncertainty band is shown in black and gray. The radii used in the clustering algorithm (anti-$k_T$) are $R = 0.4$ (left panel) and $R = 0.7$ (right panel). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 93

6.1. Sample Feynman diagram associated with giant K-factors at NLO QCD. The blob represents the hard scattering at scale $Q \gg m_W$ while the red vector boson is the subleading one which generates EW soft and collinear logarithms. . . . 96
6.2. Feynman diagrams for $W^+Z$ production at LO. Panel (a) is the $t$-channel contribution, panel (b) is the $s$-channel contribution and panels (c,d) are DY-type channels. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 97
6.3. $W^+Z$ predictions at NNLO QCD accuracy obtained using the MiNNLO$_{\mathrm{PS}}$ method (blue) and obtained at fixed-order using Matrix+OpenLoops (red) in the `inclusive setup`. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 103
6.4. $W^+Z$ predictions at NLO EW accuracy obtained using POWHEG (blue) and obtained at fixed-order using Matrix+OpenLoops (red) in the `inclusive setup`. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 104

6.5. Rapidity $y_{ee}$ (left) and invariant mass $m_{ee}$ (right) of the reconstructed $Z$ boson for $W^+Z$ production in the `inclusive setup`. . . . . . . . . . . . . . . . . . . . . 106
6.6. Missing transverse momentum $p_{\mathrm{T,miss}}$ for $W^+Z$ production in the `inclusive setup` (left) and `fiducial setup` (right). . . . . . . . . . . . . . . . . . . . . . . . . 108
6.7. Transverse momentum of the charged leading lepton $p_{\mathrm{T},\ell_1}$ for $W^+Z$ production in the `inclusive setup` (left) and `fiducial setup` (right). . . . . . . . . . . . . 109
6.8. Transverse momentum of the charged subleading lepton $p_{\mathrm{T},\ell_2}$ for $W^+Z$ production in the `inclusive setup` (left) and `fiducial setup` (right). . . . . . . . 110
6.9. Invariant mass of the three charged leptons for $W^+Z$ production in the `inclusive setup` (left) and `fiducial setup` (right). . . . . . . . . . . . . . . . . . 111
6.10. Comparison of the $\mathrm{NNLO}_{\mathrm{QCDxEW}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}$ result with MPI effects (blue and solid curve), without MPI effects (red and dashed curve) and ATLAS data. . . . . . 112

A.1. Differential distributions for $e^-\bar{\nu}_e b\bar{b}$ production with `fiducial-YR` cuts (table 4.5). See text for more details. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 122
A.2. Differential distributions for $e^-\bar{\nu}_e b\bar{b}$ production with `fiducial-YR` cuts (table 4.5). See text for more details. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 123

# List of Tables

3.1. Formal accuracy reached through POWHEG, MiNLO′ and MiNNLO$_{\text{PS}}$ for the production of a colour singlet F in association with extra QCD emissions J. . . . 35

3.2. Formal accuracy reached through the F, FJ and FJJ generators in the POWHEG framework according to the jet multiplicity. See text for more details. . . . . . 38

4.1. Inclusive cross sections for $pp \to e^+\nu_e H$ production using the MiNNLO$_{\text{PS}}$ and NNLOPS methods. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 60

4.2. Inclusive cross sections for $pp \to e^+e^- H$ production obtained through two independent MiNNLO$_{\text{PS}}$ implementations within POWHEG-BOX-RES and POWHEG-BOX-V2, and MCFM. Digits in brackets refer to statistical error only. 62

4.3. Fiducial phase space for $e^+e^- b\bar{b}$ production. . . . . . . . . . . . . . . . . . . . . 63

4.4. Fiducial cross sections for $pp \to ZH \to e^+e^- b\bar{b}$ production using the MiNNLO$_{\text{PS}}$ and NNLOPS methods. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 64

4.5. Fiducial phase space in the `fiducial-YR` and the `fiducial-ATLAS` setups. . . . 66

4.6. Integrated cross sections for $W^\pm H$ and $ZH$ production with $H \to b\bar{b}$ decay in the `inclusive` and the `fiducial-YR` setups. . . . . . . . . . . . . . . . . . . . . 67

4.7. Fiducial cross sections in the `fiducial-ATLAS` setup compared to ATLAS data. 72

5.1. Fiducial phase space used in section 5.4.3. . . . . . . . . . . . . . . . . . . . . . . . 91

6.1. Definitions of `inclusive setup` and `fiducial setup` used in the phenomenological analysis for $pp \to \mu^+\nu_\mu e^+e^-$ production. Note that leptons are always considered dressed. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 105

# 1. Introduction

The Standard Model (SM) is the unifying theory of three of the four known fundamental forces, namely the electromagnetic, weak and strong interactions. It has proven to be incredibly successful in the description of the subatomic world and, with the discovery of the long-awaited Higgs boson on the 4th of July 2012, the completion of the model was finally achieved. Our best understanding of how elementary particles interact among each other is encapsulated in the SM and its striking success is proven by the excellent ability to describe experimental data at colliders. Nowadays, the world's largest and most powerful collider is the Large Hadron Collider (LHC), in which protons collide at a centre-of-mass energy of 13 TeV and the produced particles are detected by four main experiments (ATLAS, CMS, ALICE and LHCb). The LHC was constructed as a *discovery machine* in 2008, with the main goal of detecting the last missing particle of the SM puzzle, the Higgs boson. After only a few years, the particle was observed and this milestone discovery signed the opening of a new stage: the *precision era*.

In the coming years, the LHC will be converted to a *precision machine*, which is a revolutionary step for a hadron collider. Traditionally only lepton colliders were used as precision machines, mainly because of the clean environment that allows for precise measurements. Anyway, the remarkable success of the LHC drove the physics community towards the possibility of exploiting the potential of this machine even further. In the next few years, this plan will lead to the so-called High Luminosity phase of the LHC (HL-LHC), which consists in an upgrade of detectors and an increase in the luminosity. Concretely, this translates into measurements with an unprecedented level of precision. It is thus clear that this challenging project can lead to groundbreaking results only if a significant effort is made from the theory community as well. A deep synergy between the theory and the experimental communities in this regard is crucial, as the upcoming data can be correctly interpreted only through theoretical predictions that match the experimental accuracy at least.

With the discovery of the Higgs boson, the SM has revealed to be more predictive than ever in the description of processes at colliders. Despite this, it is well known that it cannot be the final fundamental theory of nature, because it cannot describe some phenomena that are experimentally observed (e.g. neutrino oscillations, Dark Matter, Dark Energy, asymmetry in matter-antimatter) and it cannot be unified (at least, with our current understanding) with the theory of General Relativity. In this context, precision physics represents a promising path towards the discovery of physics Beyond the SM (BSM) that should appear as a discrepancy between theoretical SM predictions and experimental data. As a consequence, precise theoretical calculations are crucial to interpret accurate experimental data and to identify possible small deviations in the data/theory comparison.

At the LHC, strong interaction physics, described by *Quantum Chromodynamics* (QCD), is the dominant contribution. Increasing the accuracy of theoretical calculations concretely means including higher-order contributions in perturbative QCD. This is a valid approach in the high-energy limit because the coupling of the perturbative series (namely, the strong coupling $\alpha_s$) is small. This fundamental property of QCD is referred to as *asymptotic freedom*. The current state of the art for so-called fixed-order calculations is the inclusion of next-to-next-to leading order (NNLO) corrections, which are of order $\mathcal{O}(\alpha_s^2)$. Moreover, the increasing precision of experimental data also requires the inclusion of electroweak (EW) corrections: even though the EW coupling is much smaller than the strong one ($\alpha \ll \alpha_s$), EW corrections can have a dominant effect in some kinematic regions, especially in the deep tails of some distributions. In order to give the best description possible on the whole spectrum, the new frontier for precision predictions consists in reaching NNLO QCD and NLO EW accuracy.

Despite the enormous progress made in fixed-order computations in the past years, this approach is not sufficient for a realistic description of processes at colliders. Fixed-order computations can be performed only for few final state particles, which is in contrast to the high-multiplicity final states that are observed at detectors. Moreover, this approach fails in some kinematic regions or in presence of multiple scales due to the appearance of large logarithmic contributions. These large corrections spoil the convergence of the perturbative series. We can deal with these large contributions through resummation, which consists in reordering the perturbative series so that, when considering an appropriate range for these logarithms, the resummed series is convergent. Resummation can be performed either through analytic methods or through a parton shower: albeit the accuracy of parton showers is low, they are among the most employed tools in the high-energy physics community, thanks to their flexibility in resumming all classes of logarithmic contributions and the possibility of simulating realistic events at hadron colliders.

This thesis analyses in detail the matching between fixed-order calculations and parton showers, which is one of the key building blocks for the construction of realistic predictions for collider physics. Realistic predictions for events at colliders are usually obtained using dedicated Monte Carlo (MC) event generators that are able to model an event from the high energy limit down to the detector level. Improving the accuracy of this modelling is not an easy task, as it should account for different effects associated with different energy scales. More precisely, we can imagine an event at a collider as a sequence of phenomena that occur at different energies and regimes. This hierarchy of scales allows us to study each of these phenomena separately according to the *factorization principle* [1] for which physical effects at high energy are not affected by low-energy physics, as they are widely separated in time.

At the LHC, protons collide at high energies in the hard scattering. This scattering is called *hard* because it is related to a scale $Q$ much higher than the QCD cutoff $\Lambda_{QCD}$ where the perturbative regime breaks. This process is studied using a fixed-order computation so that we can improve the description of the hard scattering including higher-order contributions in the perturbative expansion. We should keep in mind that a complication is given by the composite nature of protons: our perturbative approach holds only for partons (quarks and gluons) in the initial state, as we cannot describe hadrons from first perturbative principles. In order to deal with this problem, we can rely again on the factorization assumption and encode the non-perturbative structure of protons into *parton distribution functions* (PDFs) that are extrapolated from data. As already discussed, sticking to a fixed-order computation is not sufficient: the hard scattering deals with a $2 \to n$ process, with $n$ much smaller than the number of final state particles observed at detectors. Moreover, a fixed-order calculation fails in some kinematic regions due to large logarithmic contributions. These issues are solved by the parton shower. The parton shower is the link between the high scale $Q$ and the detector level $\Lambda_{QCD}$. It generates subsequent emissions from the external legs of the hard scattering in the soft and collinear approximation, which captures the dominant contributions, thus producing a cascade of splittings that leads to a high-multiplicity final state. These subsequent splittings determine a degradation of the energy: the parton shower ends when the energy reaches the $\Lambda_{QCD}$ cutoff. At this stage, the perturbative regime breaks down and we need to rely on phenomenological models tuned to data for describing hadron formation. After hadronization, event generators can also describe the decay of primary hadrons to stable hadrons that reach the detectors.

It is now clear that increasing the accuracy in a realistic description of an event at a collider is extremely complicated, as it requires improving the description of different phenomena that are studied through different approaches and techniques. This thesis presents some recent developments in the matching between fixed-order computations and parton showers, as they are two complementary approaches that can be merged keeping the best features of both. The main idea of the matching procedure is to keep the high accuracy of the fixed-order computation while retaining the realistic description given by the parton shower. The current state of the art is represented by NNLO+PS accuracy, which means retaining NNLO accuracy for the hard scattering without spoiling the logarithmic structure of the parton shower. To reach this accuracy, a standard strategy is not yet available in the literature but, among the different methods at hand nowadays, the MiNNLO$_{\text{PS}}$ approach plays a crucial role because of its simplicity: it does not require a computationally demanding reweighting, as all the terms needed are already present at the level of generation of events, and it does not introduce any unphysical scale for separating the phase space according to the jet multiplicities.

In this thesis, we review some recent cutting-edge results obtained using the MiNNLO$_{\text{PS}}$ method. More precisely, this manuscript is structured as follows: in Part I we review the theoretical framework, starting from a detailed presentation of fixed-order calculations and parton shower simulations in chapter 2. We then largely discuss the problem of the matching in chapter 3: we first present the POWHEG method for obtaining NLO+PS results in section 3.1, its improvements in the MiNLO and subsequent MiNLO$'$ methods in section 3.2, and then we analyse in detail the MiNNLO$_{\text{PS}}$ approach in section 3.3.

Part II is dedicated to phenomenological studies of key processes at the LHC. In chapter 4 we present NNLO+PS results for the production of a vector boson in association with a Higgs boson ($VH$, with $V = W^{\pm}, Z$) with subsequent decay of the Higgs boson into a pair of bottom quarks ($H \to b\bar{b}$) as well as the decay of the vector boson $V$ to all possible leptonic final states. This process is fundamental for precision measurements in the Higgs sector, as it provides direct access to the bottom Yukawa coupling and it actually yields the highest sensitivity to its determination. In this context, we also discuss the differences among different clustering algorithms for the identification of $b$-jets, showing that the choice of the clustering algorithm has a non-negligible impact on differential distributions.

In chapter 5 we extend our study of $ZH$ production with $H \to b\bar{b}$ decay to the Standard Model Effective Field Theory (SMEFT). We encode in our computation a set of dimension-six operators that have a direct impact in QCD, reaching NNLO+PS accuracy in the SMEFT. We formally consider also N$^3$LO contributions that are associated to weakly constrained Wilson coefficients. The main outcome is the identification of a set of observables that are particularly suitable for increasing the sensitivity to the analysed SMEFT operators.

Chapter 6 shows results for $W^{\pm}Z$ production at NNLO QCD and NLO EW accuracy matched to parton showers. The combination of QCD and EW results is a delicate issue, especially in the presence of a parton shower: we present different possible combinations that differ for terms beyond accuracy, analysing their differences for various observables and identifying the optimal scheme. The presented method is pioneering for a possible extension of the MiNNLO$_{\text{PS}}$ framework towards the inclusion of EW effects, which is very challenging as it requires a deep understanding of the mixed QCD/EW structure of higher-order corrections within the MiNNLO$_{\text{PS}}$ formalism.

We conclude in chapter 7 where we summarize our results and we discuss possible outlooks.

This thesis is largely based on the following publications by the author:

[2] S. Zanoli, M. Chiesa, E. Re, M. Wiesemann and G. Zanderighi, *Next-to-next-to-leading order event generation for VH production with H →$b\bar{b}$ decay*, *JHEP* **07** (2022) 008, [`2112.04168`]

[3] U. Haisch, D. J. Scott, M. Wiesemann, G. Zanderighi and S. Zanoli, *NNLO event generation for $pp \to Zh \to \ell^+\ell^- b\bar{b}$ production in the SM effective field theory*, *JHEP* **07** (2022) 054, [`2204.00663`]

[4] J. M. Lindert, D. Lombardi, M. Wiesemann, G. Zanderighi and S. Zanoli, *WZ production at NNLO QCD and NLO EW matched to parton showers with MiNNLO$_{PS}$*, *JHEP* **11** (2022) 036, [`2208.12660`]

where the results described in chapters 4, 5 and 6 were originally presented.

# Part I.
# THE THEORY

# 2. Theoretical Framework

Precise collider phenomenology relies on a profound understanding of QCD over a vast range of scales, spanning from the proton mass to the centre-of-mass energy of the hard scattering. The theoretical description of a collider event is extremely complicated as an event is made up of different stages that appear at different energies: to appropriately describe each stage, various theoretical and technical tools are needed.

In this chapter, we review the two most common tools that are used for obtaining predictions at colliders in the perturbative regime: fixed-order calculations and parton showers. Our main goal is to highlight the main features of both approaches, showing that they are complementary tools that can be fully exploited when matched together: through their matching, we can obtain predictions that are both *precise* and *realistic*. In fact, fixed-order calculations deal with the hard collision and the accuracy of the calculation can be systematically improved including higher-order contributions in the perturbative expansion of the cross section. Fixed-order calculations thus allow for *precise* predictions. By contrast, parton shower simulations describe the production of a high-multiplicity final state through a cascade of splittings generated in a probabilistic way, and they thus connect the high-energy limit with the detector level. Parton showers allow for *realistic* predictions.

To understand how the two descriptions can be matched together, we now present a detailed overview of the two methods, introducing concepts and notations that will be useful in the next chapters. Fixed-order calculations are discussed in section 2.1 while parton showers are presented in section 2.2.

## 2.1. Fixed-order calculations

### 2.1.1. Preliminaries

We consider the following hadronic collision

$$h_1\, h_2 \to X\,, \tag{2.1}$$

where $h_1$ and $h_2$ are the two incoming hadrons and $X$ represents a generic final state. The incoming momenta are labelled $K_\oplus$ for $h_1$ and $K_\ominus$ for $h_2$, where the subscripts $\oplus$ and $\ominus$ refer to the direction of motion with respect to the collision axis.

Hadrons are not fundamental particles, as they present a composite structure of partons. The hard collision we are interested in involves fundamental partons that carry only a fraction of the total incoming momentum. We consider these partons aligned to the incoming hadrons and we define the longitudinal momentum fraction $x_\oplus$ for the parton extrapolated from $h_1$

and $x_\ominus$ for the parton extrapolated from $h_2$. The parton momenta $k_\oplus$ and $k_\ominus$ thus read

$$k_\oplus = x_\oplus K_\oplus \,, \qquad k_\ominus = x_\ominus K_\ominus \,, \tag{2.2}$$

with $\{x_\oplus,\, x_\ominus\} \in [0,1]$. The final state $X$ can be considered as a generic $n$-body system so that energy-momentum conservation is satisfied as follows:

$$x_\oplus K_\oplus + x_\ominus K_\ominus = k_\oplus + k_\ominus = k_1 + ... + k_n \,. \tag{2.3}$$

The cross section for this process $\sigma(h_1\, h_2 \to X)$ can be obtained relying on the factorization theorem [1] through which the high-energy (short-distance) physics is factorized from the low-energy (long-distance) one:

$$\sigma(h_1 h_2 \to X) = \sum_{i,j} \int \mathrm{d}x_\oplus \mathrm{d}x_\ominus f_{i/h_1}(x_\oplus, \mu_F^2) f_{j/h_2}(x_\ominus, \mu_F^2) \hat{\sigma}_{ij,X}(\hat{s}, \mu_R^2, \mu_F^2) + \mathcal{O}\left(\frac{\Lambda_{\mathrm{QCD}}^2}{Q^2}\right) . \tag{2.4}$$

This equation is valid up to some non-perturbative corrections encoded in the last term, where $\Lambda_{\mathrm{QCD}} \sim 0.2\,\mathrm{GeV}$ is the scale where QCD perturbation theory breaks down and $Q$ is a typical scale of the process. More details on this are given in section 2.1.2.

The term $f_{i/h_1}(x_\oplus, \mu_F^2)$ $(f_{j/h_2}(x_\ominus, \mu_F^2))$ is a parton distribution function (PDF), which represents the probability of extrapolating a parton of flavour $i$ ($j$) from $h_1$ ($h_2$) with longitudinal momentum fraction $x_\oplus$ ($x_\ominus$). PDFs are universal and they encapsulate the long-distance structure of the proton. They are obtained fitting experimental data and they are evaluated at a scale $\mu_F$ - namely the *factorization scale* - that is an unphysical scale we must introduce in our calculation for reasons that will be made clear in the next sections.

The term $\hat{\sigma}_{ij,X}$ represents the partonic cross section, which is a process-dependent quantity that encodes the high-energy dynamics. It depends on the partonic energy, defined as $\hat{s} = (k_\oplus + k_\ominus)^2$, the renormalization scale $\mu_R$ (another unphysical scale the meaning of which will be made clear later) and the factorization scale $\mu_F$. The partonic cross section is calculated as the product of an appropriate flux factor, the $n$-body phase space and the $ij \to X$ matrix element, as follows:

$$\hat{\sigma}_{ij,X}(\hat{s}, \mu_R^2, \mu_F^2) = \frac{1}{2\hat{s}} \int \prod_{l=1}^{n} \frac{\mathrm{d}^3\vec{k}_l}{(2\pi)^3 2E_l} (2\pi)^4 \delta^4\left(k_\oplus + k_\ominus - \sum_{l=1}^{n} k_l\right) |M_{ij,X}|^2 \,. \tag{2.5}$$

We note that we always assume spin and colour sums and averages already encoded in the matrix element.

The partonic cross section is a perturbative object: its accuracy can be increased including higher-order contributions in the expansion of the matrix element, which admits a perturbative series in the coupling. In fact, considering only QCD effects the matrix element can be expanded in terms of the coupling constant $g_s = \sqrt{4\pi\,\alpha_s}$ as

$$M_{ij,X} = M_{ij,X}^{(0)} + \left(\frac{g_s}{\sqrt{4\pi}}\right) M_{ij,X}^{(1)} + \left(\frac{g_s}{\sqrt{4\pi}}\right)^2 M_{ij,X}^{(2)} + \mathcal{O}\left(\left(\frac{g_s}{\sqrt{4\pi}}\right)^3\right) , \tag{2.6}$$

where the superscripts in brackets represent the perturbative order. Including contributions with increasing power of $g_s$ means accounting for terms that are increasingly suppressed, making the calculation more precise. The terms $M_{ij,X}^{(i)}$ are calculated in perturbation theory considering all the Feynman diagrams contributing to the process $ij \to X$ up to a certain power of the coupling, according to the Feynman rules of the theory. The perturbative expansion of the matrix element directly translates into an expansion of the partonic cross section. In fact, when squaring equation (2.6), we obtain

$$|M_{ij,X}|^2 = |M_{ij,X}^{(0)}|^2 + 2\alpha_s \,\mathrm{Re}[M_{ij,X}^{(0)*} \cdot M_{ij,X}^{(1)}] + \alpha_s^2 |M_{ij,X}^{(1)}|^2 + \mathcal{O}(\alpha_s^2)\,, \tag{2.7}$$

so that the partonic cross section admits the following expansion in $\alpha_s$:

$$\hat{\sigma}_{ij,X} = \hat{\sigma}_{ij,X}^{(0)} + \alpha_s \,\hat{\sigma}_{ij,X}^{(1)} + \alpha_s^2 \,\hat{\sigma}_{ij,X}^{(2)} + \mathcal{O}(\alpha_s^3)\,. \tag{2.8}$$

Considering only the first order proportional to $\alpha_s^0$, we are performing a *leading order* (LO) calculation. Adding one power of the coupling $\alpha_s$ leads to a *next-to-leading* (NLO) calculation. With two powers of the coupling $\alpha_s^2$ we have a *next-to-next-to-leading order* (NNLO) calculation, and so on. We stress again that here we consider only QCD corrections because $\alpha_s$ is much larger than the electroweak coupling $\alpha$. However, given the unprecedented level of accuracy required nowadays for precision phenomenology, the inclusion of subleading EW effects in precise theoretical predictions is becoming urgent. Including EW effects consists not only in including contributions proportional to the EW coupling but also mixed QCD/EW corrections $\mathcal{O}(\alpha_s\,\alpha)$. These calculations are highly non-trivial and they represent the current frontier for precision physics. We will discuss in detail this problem in chapter 6.

To make the discussion easier, we now introduce the notation used in [5] and we will stick to it for the rest of this work. The total cross section for $h_1\, h_2 \to X\,(= k_1 + \ldots + k_n)$ at LO is

$$\sigma_{\mathrm{LO}} = \int \mathrm{d}\boldsymbol{\Phi}_n \mathcal{L}\, \mathcal{B}(\boldsymbol{\Phi}_n)\,, \tag{2.9}$$

where $\mathrm{d}\boldsymbol{\Phi}_n$ includes the differential parton energy fractions ($\mathrm{d}x_\oplus$, $\mathrm{d}x_\ominus$) and the differential $n$-body phase space ($\mathrm{d}\Phi_n$)

$$\mathrm{d}\boldsymbol{\Phi}_n = \mathrm{d}x_\oplus \mathrm{d}x_\ominus \mathrm{d}\Phi_n = \mathrm{d}x_\oplus \mathrm{d}x_\ominus \prod_{l=1}^{n} \frac{\mathrm{d}^3\vec{k}_l}{(2\pi)^3 2E_l} (2\pi)^4 \delta^4\left(k_\oplus + k_\ominus - \sum_{l=1}^{n} k_l\right)\,, \tag{2.10}$$

and the product $\mathcal{L}\,\mathcal{B}(\boldsymbol{\Phi}_n)$ implicitly assumes a sum over the incoming flavours:

$$\mathcal{L}\,\mathcal{B}(\boldsymbol{\Phi}_n) = \sum_{ij} \mathcal{L}_{ij}\, \mathcal{B}_{ij}(\boldsymbol{\Phi}_n)\,. \tag{2.11}$$

The luminosity factor $\mathcal{L}$ is defined as the product of the PDFs

$$\mathcal{L} = \mathcal{L}_{ij}(x_\oplus, x_\ominus) = f_{i/h_1}(x_\oplus, \mu_F^2) f_{j/h_2}(x_\ominus, \mu_F^2)\,, \tag{2.12}$$

and $\mathcal{B}(\boldsymbol{\Phi}_n)$, commonly referred to as *Born* contribution, is the LO squared matrix element weighted with the appropriate flux factor:

$$\mathcal{B}(\boldsymbol{\Phi}_n) = \mathcal{B}_{ij}(\boldsymbol{\Phi}_n) = \frac{|M^{(0)}_{ij,X}(\boldsymbol{\Phi}_n)|^2}{4k_\oplus k_\ominus} \,. \tag{2.13}$$

When performing a NLO computation, we need to encode both real and virtual corrections: real contributions are associated to the production of an extra particle in the final state while virtual corrections arise from the interference of one-loop diagrams with the Born level. The NLO cross section can be written as follows

$$\sigma_{\mathrm{NLO}} = \int \mathrm{d}\boldsymbol{\Phi}_n \mathcal{L} \Big[ \mathcal{B}(\boldsymbol{\Phi}_n) + \mathcal{V}(\boldsymbol{\Phi}_n) \Big] + \int \mathrm{d}\boldsymbol{\Phi}_{n+1} \mathcal{L} R(\boldsymbol{\Phi}_{n+1}) \,, \tag{2.14}$$

where $\mathcal{V}(\boldsymbol{\Phi}_n)$ and $R(\boldsymbol{\Phi}_{n+1})$ are the virtual and the real contributions, respectively. The $(n+1)$-body phase space is straightforwardly given by:

$$\mathrm{d}\boldsymbol{\Phi}_{n+1} = \mathrm{d}x_\oplus \mathrm{d}x_\ominus \mathrm{d}\Phi_{n+1} = \mathrm{d}x_\oplus \mathrm{d}x_\ominus \prod_{l=1}^{n+1} \frac{\mathrm{d}^3\vec{k}_l}{(2\pi)^3 2E_l} (2\pi)^4 \delta^4\Big( k_\oplus + k_\ominus - \sum_{l=1}^{n+1} k_l \Big) \,. \tag{2.15}$$

The calculation of $\sigma_{\mathrm{NLO}}$ in (2.14) is non-trivial as both virtual and real corrections are affected by singularities. In general, while performing calculations at NLO (and higher) accuracy, cross sections are affected by ultraviolet (UV) and infrared (IR) divergences. UV divergences arise when integrating over loop momenta while IR singularities are associated with soft and/or collinear emissions. These singularities thus have a different origin and are handled through different theoretical and technical tools. In order to fully understand the structure of a fixed-order computation, we will review in detail how to remove UV and IR divergences in the next sections, 2.1.2 and 2.1.3 respectively.

### 2.1.2. Ultraviolet divergences

When calculating loop integrals in virtual contributions, the momentum $\ell$ flowing in the loop is not constrained by momentum conservation and should be integrated up to infinity. In general, loop integrals have the following form

$$\int_0^\infty \frac{\mathrm{d}^4\ell}{(2\pi)^4} \frac{1}{(\ell^2 - m^2)^k} \,, \tag{2.16}$$

where the power $k$ in the denominator depends on the specific propagators we are considering. The integral can present singularities in the limit $\ell \to \infty$ (if $2k - 4 \leq 0$), and thus these singularities are refferred to as ultraviolet (UV) divergences.

In Quantum Field Theory, UV divergences are handled through a well-established procedure. Firstly, divergent integrals are subject to *regularization*, which consists in isolating the divergences in the calculation. More precisely, we introduce extra parameters in the computation so that the result depends on these parameters and is divergent in the corresponding physical limit. The most common options are [6]:

- Cut-off regularization: the range of the loop integral is constrained by a cut-off scale $\Lambda$. The physical limit is restored if $\Lambda \to \infty$.

$$\int_0^\infty \frac{\mathrm{d}^4\ell}{(2\pi)^4} \frac{1}{(\ell^2 - m^2)^k} \quad \to \quad \int_0^\Lambda \frac{\mathrm{d}^4\ell}{(2\pi)^4} \frac{1}{(\ell^2 - m^2)^k} \, . \tag{2.17}$$

  Notice that this method is not Lorentz invariant, as the parameter $\Lambda$ depends on the frame of reference.

- Pauli-Villars regularization: auxiliary fields of mass $M$ are added to the Lagrangian, or directly as mass terms in the propagators, in order to render the integral convergent. The physical limit is restored when $M \to \infty$, as the propagators of these extra particles vanish.

$$\int_0^\infty \frac{\mathrm{d}^4\ell}{(2\pi)^4} \frac{1}{(\ell^2 - m^2)^k} \quad \to \quad \int_0^\infty \frac{\mathrm{d}^4\ell}{(2\pi)^4} \left[ \frac{1}{(\ell^2 - m^2)^k} - \frac{1}{(\ell^2 - M^2)^k} \right] . \tag{2.18}$$

- Dimensional regularization: the dimensionality of the phase space is conventionally reduced from $d = 4$ to $d = 4 - 2\epsilon$, with $\epsilon > 0$. The physical limit corresponds to $\epsilon \to 0$ and singularities appear as poles in $\epsilon$.

$$\int_0^\infty \frac{\mathrm{d}^4\ell}{(2\pi)^4} \frac{1}{(\ell^2 - m^2)^k} \quad \to \quad \int_0^\infty \frac{\mathrm{d}^{4-2\epsilon}\ell}{(2\pi)^{4-2\epsilon}} \frac{1}{(\ell^2 - m^2)^k} \, . \tag{2.19}$$

  In this procedure, the dimensionality of all the objects of the theory changes accordingly.

Once loop integrals are regularized, the result depends explicitly on the regularization parameters ($\Lambda$, $M$ or $\epsilon$ in the methods presented above) and divergences can be restored in the corresponding physical limit.

The second step needed for handling UV divergences is called *renormalization*. Renormalization is a procedure that eliminates UV divergences expressing unphysical (or *bare*) quantities in terms of physical parameters. The idea behind the renormalization procedure is the following: we start from a bare Lagrangian that contains bare parameters. These bare parameters are not physical, in the sense that they are not measurable in laboratories. At this stage, loop integrals are divergent in the UV limit. After regularization, the isolated divergences can be reabsorbed into a redefinition of the parameters of the theory that now become *physical* parameters. Concretely, renormalization consists in adding a set of counterterms in the Lagrangian that compensate for the UV divergences order by order in perturbation theory. These counterterms are associated with renormalization constants $Z$ that relate physical and bare quantities together. Counterterms can be constructed using different *renormalization schemes*, which are related to different possibilities in the choice of finite parts. One of the most commonly used is the modified minimal subtraction scheme ($\overline{\text{MS}}$) in which poles are subtracted together with the term $\log(4\pi) - \gamma_{\mathrm{E}}$, where $\gamma_{\mathrm{E}}$ is the Eulero-Mascheroni constant. This choice is particularly convenient because the finite expression that is removed always arises in the UV limit of Feynman integrals.

Once the Lagrangian is expressed only in terms of physical objects, divergences are not present anymore, as they are hidden in the definition of the new renormalized parameters. In this procedure, physical parameters gain a dependence on a scale, called *renormalization scale* ($\mu_R$): it represents an unphysical scale needed to perform the calculation, but physics does not change when varying it.

The SM is a *renormalizable* theory, which means that the number of counterterms needed in the renormalization procedure is finite and that this strategy can be applied to all orders in perturbation theory. Thus, the SM is free from UV divergences. Considering the specific case of QCD, the renormalization of the coupling constant is particularly important as it is related to the fundamental property of *asymptotic freedom*. The bare coupling constant $\alpha_s^{\text{bare}}$ is related to the renormalized one $\alpha_s^{\text{ren}}$ through the renormalization constant $Z_{\alpha_s}$ as follows (see standard textbooks [6, 7])

$$\alpha_s^{\text{bare}} = \mu_R^{2\epsilon}\, Z_{\alpha_s}^2\, \alpha_s^{\text{ren}}\,, \tag{2.20}$$

where $\epsilon$ is the regularization parameter in dimensional regularization. The renormalization constant $Z_{\alpha_s}$ is determined order by order in perturbation theory with the following expansion

$$Z_{\alpha_s} = 1 + \delta Z_{\alpha_s}^{(1)}\,. \tag{2.21}$$

Bare quantities do not have any dependence on the renormalization scale: if we take the derivative of (2.20) with respect to $\mu_R$, we obtain an equation that governs the running of the renormalized strong coupling when varying the renormalization scale:

$$\frac{\mathrm{d}\alpha_s^{\text{bare}}}{\mathrm{d}\mu_R^2} = 0 \qquad \rightarrow \qquad \frac{\partial \alpha_s(\mu_R)}{\partial \mu_R^2} = -\alpha_s(\mu_R)\frac{1}{\mu_R^2 Z_{\alpha_s}^2}\frac{\partial \mu_R^{2\epsilon} Z_{\alpha_s}^2}{\partial \mu_R^2}\,. \tag{2.22}$$

Note that we dropped the superscript in the renormalized couplings. Unless explicitly stated, from now on all the parameters are considered renormalized. In the limit $\epsilon \to 0$, equation (2.22) can be rewritten as

$$\mu_R^2 \frac{\partial \alpha_s(\mu_R)}{\partial \mu_R^2} \equiv \beta(\alpha_s) = -\alpha_s(\mu_R)\sum_{k=1}^{\infty} \beta_{k-1}(\alpha_s)\alpha_s^k(\mu_R)\,, \tag{2.23}$$

which is the *renormalization group equation* (RGE) for the running coupling. It describes how the renormalized strong coupling varies according to the renormalization scale. This dependence is governed by the QCD $\beta$-function that shows a perturbative expansion starting from $\mathcal{O}(\alpha_s^2)$:

$$\beta(\alpha_s) = -\beta_0\, \alpha_s^2 - \beta_1\, \alpha_s^3 - \beta_2\, \alpha_s^4 + \mathcal{O}(\alpha_s^5)\,. \tag{2.24}$$

The $\beta_k$ coefficients are calculated in perturbation theory and, considering the convention used in [8], they read

$$\beta_0 = \frac{11C_{\text{A}} - 2n_f}{12\pi}\,, \tag{2.25}$$

$$\beta_1 = \frac{17C_{\text{A}}^2 - 5C_{\text{A}}n_f - 3C_{\text{F}}n_f}{24\pi^2}\,, \tag{2.26}$$

$$\beta_2 = \frac{2857C_{\text{A}}^3 + (54C_{\text{F}}^2 - 615C_{\text{A}}C_{\text{F}} - 1415C_{\text{A}}^2)n_f + (66C_{\text{F}} + 79C_{\text{A}})n_f^2}{3456\pi^3}\,, \tag{2.27}$$

where $C_\mathrm{A}$ and $C_F$ are the colour factors, with values 3 and 4/3 respectively, and $n_f$ is the number of active flavours. Notice that $\beta_2$ explicitly depends on the scheme used. The expression presented above is in the $\overline{\mathrm{MS}}$ scheme.

Including in the $\beta$-function only the first coefficient $\beta_0$, we can solve (2.23) analytically:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1+\beta_0 \ln\left(Q^2/\mu^2\right)\alpha_s(\mu^2)}\,. \tag{2.28}$$

The previous equation relates the value of the strong coupling at different scales: once $\alpha_s$ is known at a scale $\mu$, we can obtain its value at any other scale $Q$. Even though we can use perturbation theory to evolve the strong coupling at any scale, its absolute value is not known and it has to be obtained from experiments. Nowadays, the most common choice is using as input the value of $\alpha_s$ at the mass of the $Z$ boson, $\alpha_s(m_Z^2) \sim 0.118$.

Given equation (2.28) and the negative sign of the $\beta$-function, it is clear that the strong coupling becomes smaller when the energy increases:

$$\alpha_s(Q^2) \to 0 \quad \text{when} \quad Q \to \infty\,. \tag{2.29}$$

This property is the asymptotic freedom of QCD: at high energies, quarks and gluons interact weakly and they can be considered as free particles. This property is crucial as it allows us to use perturbation theory for obtaining predictions in QCD. By contrast, at low energies perturbation theory breaks down because the coupling constant gets larger and larger. We can obtain an estimate of where perturbation theory is no longer allowed and non-perturbative effects become dominant studying where the coupling diverges. Imposing that the denominator in (2.28) is zero, we find

$$\alpha_s(Q^2) = \frac{1}{\beta_0 \ln\left(Q^2/\Lambda^2_{QCD}\right)}\,, \tag{2.30}$$

where $\Lambda_\mathrm{QCD}$ is the energy scale at which the coupling becomes singular. $\Lambda_\mathrm{QCD}$ is around 220 MeV [9] and, below this scale, perturbation theory fails.

At this point the presence of the renormalization scale in equation (2.4) is clear: the partonic cross section $\hat{\sigma}_{ij}$ depends explicitly on renormalized parameters and thus on the renormalization scale $\mu_R$. We remind the reader that this scale is unphysical: if we were able to compute the hadronic cross section to all orders in perturbation theory, the dependence on $\mu_R$ would disappear. Since this is not feasible, the cross section is concretely obtained up to a certain accuracy truncating its perturbative expansion at a certain order in the coupling, generating a residual dependence on the renormalization scale.

In order to avoid potentially large logarithmic contributions $\log\left(Q^2/\mu^2\right)$ that can spoil the convergence of the perturbative series, the renormalization scale is usually set to a typical scale $Q$ of the process under study. When dealing with complex processes, multiple scales can be present and the choice of $\mu_R$ is not trivial. In these cases, there is no correct choice a priori. We will discuss this problem in section 3.2.

### 2.1.3. Infrared divergences

When calculating higher-order QCD corrections, UV divergences are not the only kind of singularities appearing. In fact, also divergences associated with soft and/or collinear emissions emerge and they are thus called *infrared* (IR) singularities. Before analysing in detail how these singular behaviours appear in perturbation theory, we start our discussion from the Kinoshita-Lee-Nauenberg (KLN) theorem [10, 11], which guarantees their cancellation in inclusive calculations.

The KLN theorem states that soft and collinear singularities cancel in the sum over degenerate states. Examples of degenerate states are real emissions in which a quark emits a soft gluon or a quark emits a collinear gluon and a virtual correction to a quark line. These states cannot be distinguished from each other, as soft or collinear or virtual gluons cannot be detected. More precisely, the KLN theorem guarantees that real and virtual corrections show the same singularities (with opposite sign) in the soft and/or collinear limit, so that the computation of sufficiently inclusive observables is free from singularities. We stress that this theorem applies to all orders in perturbation theory. The SM is free of IR divergences, provided configurations involving all possible final and initial states are considered.

In order to analyse how IR divergences arise in NLO QCD computation, we can consider the production of a $q\bar{q}$ system from $e^+e^-$ annihilation

$$e^+(k_1)\,e^-(k_2) \to q(p_1)\,\bar{q}(p_2)\,, \tag{2.31}$$

which represents a nice framework for introducing the topic of IR singularities because the initial state is not affected by QCD effects. In fact, initial state QCD radiation deserves a specific treatment that will be discussed later in this section.

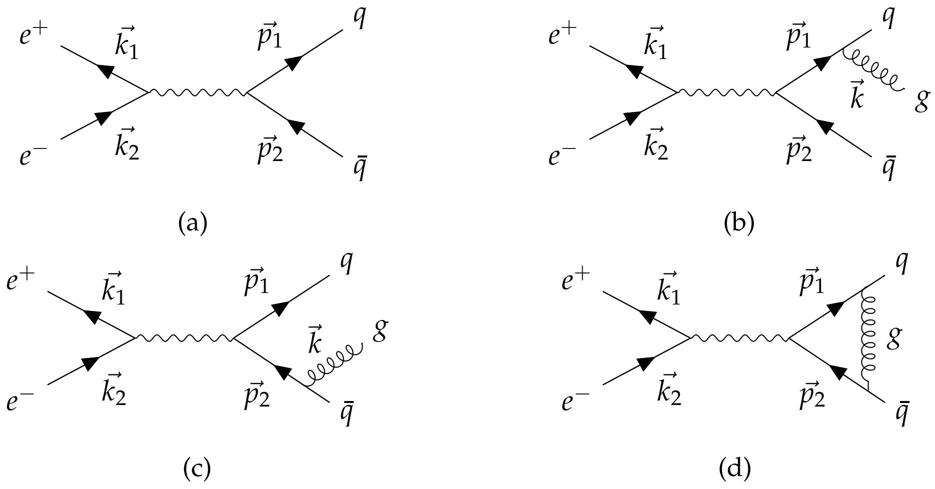


Figure 2.1.: Feynman diagrams contributing up to NLO accuracy in the process $e^+e^- \to q\bar{q}$. Panel (a) shows the LO diagram, panels (b) and (c) are the real corrections and (d) is the virtual contribution.

Sample Feynman diagrams of this process are shown in figure 2.1. The calculation of the cross section at LO and NLO is trivial, so we report here only the main results. The interested reader is referred to [12] or standard textbooks like [7].

At LO accuracy, the matrix element reads

$$M_{q\bar{q}} = \frac{e^2 Q_q}{q^2}[\bar{u}(p_1)\gamma_\mu v(p_2)][\bar{v}(k_2)\gamma^\mu u(k_1)]\,, \tag{2.32}$$

where $e$ is the electric charge, $Q_q$ is the fractional electric charge of the quarks and $q^2 = (k_1 + k_2)^2 = (p_1 + p_2)^2$. The 2-body phase space and flux factors are

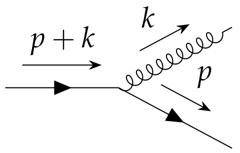
$$\mathrm{d}\Phi_2 = \frac{1}{32\pi^2}\sin\theta \mathrm{d}\theta \mathrm{d}\phi\,, \qquad \text{flux} = \frac{1}{2q^2}\,, \tag{2.33}$$

so that the LO cross section can be easily obtained as:

$$\sigma_{q\bar{q}} = \int \mathrm{d}\Phi_2 \frac{|M_{q\bar{q}}|^2}{\text{flux}} = \frac{4}{3}\alpha^2\,\frac{\pi N_c Q_q^2}{q^2}\,. \tag{2.34}$$

$\alpha = e^2/(4\pi)$ is the fine-structure constant and $N_c = 3$ is the number of colours.

The calculation of the NLO cross section consists in including both real and virtual corrections. We can already expect the appearance of singularities in the real matrix element. In fact, if we consider the emission of a gluon from a quark:



$$\text{with} \quad \begin{aligned} p &= E_p(1,0,0,1)\,, \\ k &= E_k(1,0,\sin\theta,\cos\theta)\,, \end{aligned} \tag{2.35}$$

the internal propagator is proportional to

$$\frac{1}{(p+k)^2} = \frac{1}{2E_p E_k(1-\cos\theta)}\,, \tag{2.36}$$

which shows a singularity when the gluon becomes *soft* ($E_k \to 0$) or *collinear* ($\theta \to 0$). We note that if we consider a massive quark, the collinear singularity disappears, as the propagator scales like:

$$\frac{1}{(p+k)^2 - m^2} = \frac{1}{2E_p E_k(1-\beta\cos\theta)} \quad \text{with} \quad \beta = \sqrt{1 - m^2/E^2}\,. \tag{2.37}$$

For this reason, collinear singularities are often called *mass singularities*.

The real matrix element reads

$$\begin{aligned} M_{q\bar{q}g} = &\left[\bar{u}(p_1)(-ig_s t^{\mathrm{A}} \not\epsilon)\frac{i(\not p_1 + \not k)}{(p_1+k)^2}(-ie\gamma^\mu)v(p_2)\right]\frac{(e^2Q_q)^2}{q^2}\left[\bar{v}(k_2)\gamma_\mu u(k_1)\right] \\ &+ \left[\bar{u}(p_1)(-ie\gamma^\mu)\frac{-i(\not p_2 + \not k)}{(p_2+k)^2}(-ig_s t^{\mathrm{A}}\not\epsilon)v(p_2)\right]\frac{(e^2Q_q)^2}{q^2}\left[\bar{v}(k_2)\gamma_\mu u(k_1)\right]\,, \end{aligned} \tag{2.38}$$

where the two lines correspond to diagrams 2.1(b) and 2.1(c), respectively. $k$ and $\epsilon$ are the momentum and the polarization vector of the emitted gluon and $t^{\mathrm{A}}$ is the SU(3) generator.

**Soft limit**

If we consider the soft limit ($k \to 0$), the real matrix element becomes

$$M_{q\bar{q}g} = -ieg_s t^{\mathrm{A}} \left[ \bar{u}(p_1)\gamma^\mu \left( \frac{\not\epsilon \not p_1}{2p_1 k} - \frac{\not\epsilon \not p_2}{2p_2 k} \right) v(p_2) \right] \frac{(e^2 Q_q)^2}{q^2} \left[ \bar{v}(k_2)\gamma_\mu u(k_1) \right] \tag{2.39}$$

and, squaring it, we observe the following factorization

$$R = |M_{q\bar{q}g}|^2 \xrightarrow{soft} |M_{q\bar{q}}|^2 \cdot C_{\mathrm{F}}\, g_s^2 \,\frac{2p_1 p_2}{(p_1 k)(p_2 k)} = \mathcal{B} \cdot C_{\mathrm{F}}\, g_s^2 \,\frac{2p_1 p_2}{(p_1 k)(p_2 k)}\,, \tag{2.40}$$

where $C_{\mathrm{F}} = 4/3$ is the QCD color factor. In the soft limit, the real contribution factorizes and can be written as the product of the Born contribution $\mathcal{B}$ with an appropriate Eikonal factor. A similar factorization holds also for the phase space. We can write the 3-body phase space in terms of the 2-body one times a phase space for the radiation, as follows:

$$\mathrm{d}\Phi_{q\bar{q}g} \xrightarrow{soft} \mathrm{d}\Phi_{q\bar{q}} \cdot \mathrm{d}\Phi_{rad} = \mathrm{d}\Phi_{q\bar{q}} \cdot \frac{\mathrm{d}^3\vec{k}}{(2\pi)^3 2E_k}\,. \tag{2.41}$$

This allows us to write the cross section in the soft limit in a factorized form:

$$\sigma_{q\bar{q}g} \xrightarrow{soft} \int \mathrm{d}\Phi_{q\bar{q}}\, \mathcal{B} \cdot \int \frac{\mathrm{d}^3\vec{k}}{(2\pi)^3 2E_k}\, C_{\mathrm{F}}\, g_s^2 \,\frac{2p_1 p_2}{(p_1 k)(p_2 k)} = \sigma_{q\bar{q}} \cdot \mathcal{S}\,. \tag{2.42}$$

The universal $\mathcal{S}$ function describes the emission of a soft gluon.

**Collinear limit**

The collinear limit of the real cross section is better understood if we use the Sudakov parametrization

$$p_1^\mu = z p^\mu + k_\perp^\mu - \frac{k_\perp^2}{z}\frac{n^\mu}{2p\cdot n}\,, \tag{2.43}$$

$$k^\mu = (1-z) p^\mu - k_\perp^\mu - \frac{k_\perp^2}{1-z}\frac{n^\mu}{2p\cdot n}\,, \tag{2.44}$$

where $p^\mu$ is a light-like vector defining the collinear direction and $n^\mu$ is a light-like auxiliary vector. $k_\perp^\mu$ is a vector with only two non-zero components given by the transverse momentum of the gluon. Both $p^\mu$ and $n^\mu$ are orthogonal to $k_\perp^\mu$, i.e. $k_\perp \cdot p = k_\perp \cdot n = 0$. The splitting energy $z$ is constructed as $z = E_1/(E_1 + E_k)$, which defines how the energy is shared between the quark and the gluon in the splitting. With this parametrization, the propagator in (2.38) becomes

$$\frac{1}{2p_1 \cdot k} = -\frac{z(1-z)}{k_\perp^2} \tag{2.45}$$

and the collinear limit corresponds to $k_\perp \to 0$.

The real squared matrix element can be rewritten as

$$R = |M_{q\bar{q}g}|^2 \xrightarrow{coll} |M_{q\bar{q}}|^2 \cdot C_F\, g_s^2 \,\frac{1}{p_1 \cdot k}\,\frac{1+z^2}{1-z} = \mathcal{B} \cdot C_F\, g_s^2 \,\frac{1}{p_1 \cdot k}\,\frac{1+z^2}{1-z}\,, \tag{2.46}$$

where the Born contribution $\mathcal{B}$ is obtained clustering together the two particles involved in the splitting, reconstructing the emitter quark with momentum equal to the sum of the quark ($p_1$) and gluon ($k$) momenta. The phase space factorizes as well in the collinear limit as

$$\mathrm{d}\Phi_{q\bar{q}g} \xrightarrow{coll} \mathrm{d}\Phi_{q\bar{q}} \cdot \mathrm{d}\Phi_{rad} = \mathrm{d}\Phi_{q\bar{q}}\,\frac{1}{(4\pi)^2}\left(-\frac{1}{z}\right)\frac{\mathrm{d}z}{(1-z)}\,\mathrm{d}k_\perp^2\,, \tag{2.47}$$

so that the cross section can be factorized as follows:

$$\sigma_{q\bar{q}g} \xrightarrow{coll} \int \mathrm{d}\Phi_{q\bar{q}}\,\mathcal{B} \cdot \frac{\alpha_s}{2\pi}\int \frac{\mathrm{d}k_\perp^2}{k_\perp^2}\,\mathrm{d}z\,C_F\frac{1+z^2}{1-z} = \sigma_{q\bar{q}} \cdot \mathcal{C}. \tag{2.48}$$

The universal $\mathcal{C}$ function describes the emission of a collinear gluon.

The KLN theorem [10, 11] guarantees that soft and collinear divergences cancel when summing real and virtual contributions and we thus expect virtual corrections to have the same singular behaviour (with opposite sign) as the real term in the IR regions. In fact, virtual corrections are divergent when the momentum flowing in the loop $\ell$ becomes soft, $\ell \to 0$, or collinear to an external momentum $k$, $\ell \cdot k = 0$.

To see explicitly the singular behaviour of the virtual contribution for the case of $e^+e^- \to q\bar{q}$, we regularize the loop integral through dimensional regularization: the dimensionality of the phase space is customary increased from $d = 4$ to $d = 4 + 2\epsilon$, with $\epsilon > 0$, and the physical limit is given by $\epsilon \to 0$. Also in this case, the calculation of real and virtual contributions in dimensional regularization is straightforward and not very informative, so we highlight here only the main results. The full calculation is presented in [12].

In dimensional regularization, real and virtual cross sections read

$$\sigma_{q\bar{q}g} = \sigma_{q\bar{q}} \cdot \frac{\alpha_s(\mu_R)}{2\pi}\,C_F\,\frac{\Gamma^2(1-\epsilon)}{\Gamma(1-3\epsilon)}\left(\frac{4\pi\mu_R^2}{q^2}\right)\left\{\frac{2}{\epsilon^2}+\frac{3}{\epsilon}+\frac{19}{2}+\mathcal{O}(\epsilon)\right\}, \tag{2.49}$$

$$\sigma_{q\bar{q}}^{virt} = \sigma_{q\bar{q}} \cdot \frac{\alpha_s(\mu_R)}{2\pi}C_F\,\frac{\Gamma^2(1-\epsilon)}{\Gamma(1-3\epsilon)}\left(\frac{4\pi\mu_R^2}{q^2}\right)\left\{-\frac{2}{\epsilon^2}-\frac{3}{\epsilon}-8+\mathcal{O}(\epsilon)\right\}, \tag{2.50}$$

where we explicitly see the appearance of poles in $\epsilon$. Single poles represent configurations in which the gluon is either soft or collinear, while double poles are associated to soft and collinear topologies. The coefficients of the poles in the real and virtual cross sections are equal in value but with opposite sign so that the final cross section is finite.

The explicit cancellation of soft and collinear divergences that was presented for the simple case of $e^+e^- \to q\bar{q}$ is a general property of QCD and it is satisfied to all orders in perturbation theory. In fact, the soft and collinear limits of the amplitudes are universal. If we consider a general $2 \to n$ process, they read [13]:

- Soft limit:

  $$|M_{n+1}(p_1,...p_i,...p_{n+1})|^2 \xrightarrow{p_i\to 0} \sum_{\ell,r}^{n} \mathcal{S}_{\ell r}(p_i)|M_n(p_1,...,p_n)|^2\,, \tag{2.51}$$

  where the universal Eikonal factor $\mathcal{S}_{\ell r}(p_i)$ is proportional to:

  $$\mathcal{S}_{\ell r}(p_i) \sim \frac{2p_\ell \cdot p_r}{(p_\ell \cdot p_i)(p_r \cdot p_i)}\,. \tag{2.52}$$

  In the soft limit, the real correction is equal to the Born contribution, in which the soft parton is removed, multiplied by a universal term that encodes the emission of the soft particle.

- Collinear limit:

  $$|M_{n+1}(p_1,...p_i,p_j...p_{n+1})|^2 \xrightarrow{p_i\|p_j} \mathcal{C}(z)|M_n(p_1,...,p_i+p_j,...,p_n)|^2\,, \tag{2.53}$$

  where the universal collinear function $\mathcal{C}(z)$ is proportional to

  $$\mathcal{C}(z) \sim \hat{P}_{k\to ij}(z)\,. \tag{2.54}$$

  The parton $k$ is the mother particle that splits into the $i$ and $j$ daughters. $z$ represents the energy fraction that the parton $i$ carries away from parton $k$ in the $k \to ij$ splitting. The universal $\hat{P}_{k\to ij}$ function is the *unregularized Altarelli-Parisi splitting kernel*, which depends only on the flavour of the partons undergoing the $k \to ij$ splitting. These kernels have a perturbative expansion in the strong coupling and at LO they read:

  $$\begin{aligned}
  \hat{P}^{(0)}_{qg}(z) &= \frac{1}{2}\left[z^2 + (1-z)^2\right],\\
  \hat{P}^{(0)}_{qq}(z) &= C_\mathrm{F}\left[\frac{1+z^2}{1-z}\right],\\
  \hat{P}^{(0)}_{gq}(z) &= C_\mathrm{F}\left[\frac{1+(1-z)^2}{z}\right],\\
  \hat{P}^{(0)}_{gg}(z) &= 2\,C_\mathrm{A}\left[\frac{z}{1-z} + \frac{1-z}{z} + z(1-z)\right].
  \end{aligned} \tag{2.55}$$

  The real correction with the parton $i$ becoming collinear to the parton $j$ is equal to the Born contribution in which the two collinear partons are clustered together into parton $k$ ($p_k = p_i + p_j$) multiplied by a universal splitting factor that depends only on the nature of the $k \to ij$ splitting.

These poles in the IR regions cancel when summing together real and virtual contributions thanks to the KLN theorem. This cancellation has a specific physical meaning: the cross section is a physical quantity, so it must be finite. If the real contribution is divergent, the virtual one must present the same singularity with opposite sign for conservation of probability, referred to as *unitarity*.

We can now generalise equation (2.14) to obtain the expectation value of a general observable $\mathcal{O}$ at NLO accuracy

$$\langle \mathcal{O} \rangle = \int \mathrm{d}\mathbf{\Phi}_n \mathcal{L}\, \mathcal{O}_n(\mathbf{\Phi}_n) \Big[ \mathcal{B}(\mathbf{\Phi}_n) + \mathcal{V}(\mathbf{\Phi}_n) \Big] + \int \mathrm{d}\mathbf{\Phi}_{n+1} \mathcal{L}\, \mathcal{O}_{n+1}(\mathbf{\Phi}_{n+1}) R(\mathbf{\Phi}_{n+1}) \,, \tag{2.56}$$

where $\mathcal{O}_n$ and $\mathcal{O}_{n+1}$ are the expressions of the observable $\mathcal{O}$ in terms of the $n$ and $n+1$ configurations, respectively. We remind the reader that the KLN theorem ensures that IR divergences cancel when considering all possible initial and final configurations: if we want to be more exclusive, this cancellation does not fully apply. In general, equation (2.56) is free from IR singularities if $\mathcal{O}$ is *infrared (and collinear) safe*, which means that it is not sensitive to long-physics effects. An IR (and collinear)-safe observable satisfies

$$\begin{aligned} \mathcal{O}_{n+1}(p_1, ..., p_j, ...p_{n+1}) &\xrightarrow{p_j \to 0} \mathcal{O}_n(p_1, ..., p_{j-1}, p_{j+1}...p_n) \,, \\ \mathcal{O}_{n+1}(p_1, ..., p_j, p_\ell, ...p_{n+1}) &\xrightarrow{p_j \| p_\ell} \mathcal{O}_n(p_1, ..., p_j + p_\ell, ...p_n) \,, \end{aligned} \tag{2.57}$$

which means that $\mathcal{O}$ is insensitive to extra soft and/or collinear emissions. This property is crucial because it allows us to use fixed-order calculations in perturbative QCD for obtaining predictive results.

When calculating explicitly the IR behaviour of the cross section at NLO accuracy, we considered the specific case of $e^+e^-$ annihilation, for which QCD radiation does not affect the initial state. When dealing with QCD emissions in the initial state, the KLN theorem does not fully apply. More precisely, collinear singularities do not cancel because the emission of a gluon in the collinear limit modifies the momentum of the parton undergoing the hard scattering. An illustrative example is figure (2.2).

The non-complete cancellation of collinear singularities for initial-state emissions is handled through *factorization*. Factorization is a procedure similar to renormalization but applied to singularities that arise in the collinear limit for initial-state radiation. Bare divergent PDFs entering the calculation $f^{\text{bare}}(x)$ can be reabsorbed into finite physical PDFs $f(x, \mu_F^2)$ that now gain an explicit dependence on a new unphysical scale called *factorization scale* ($\mu_F$). This scale appears in the theory in an analogous way to the renormalization one. Since the collinear behaviour of the cross section is universal, the factorization procedure can be handled in a process-independent way.

We recall that PDFs are non-perturbative objects and their values at a certain scale $\mu_F$ and for a specific momentum fraction $x$ cannot be calculated in perturbation theory but must be fitted to data. However, the evolution of the PDFs with $\mu_F$ is perturbative and it is described by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equation [14, 15, 16]. The DGLAP equation is a $(2n_f + 1)$-dimensional equation in flavour space of the form

$$\frac{\partial}{\partial \ln \mu^2} \begin{pmatrix} f_{q_i/h}(x, \mu^2) \\ f_{g/h}(x, \mu^2) \end{pmatrix} = \frac{\alpha_s(\mu^2)}{2\pi} \sum_{q_j, \bar{q}_j} \int_x^1 \frac{\mathrm{d}z}{z} \begin{bmatrix} P_{q_i q_j}\left(\frac{x}{z}, \alpha_s(\mu^2)\right) & P_{q_i g}\left(\frac{x}{z}, \alpha_s(\mu^2)\right. \\ P_{g q_j}\left(\frac{x}{z}, \alpha_s(\mu^2)\right) & P_{gg}\left(\frac{x}{z}, \alpha_s(\mu^2)\right. \end{bmatrix} \begin{pmatrix} f_{q_j/h}(z, \mu^2) \\ f_{g/h}(z, \mu^2) \end{pmatrix} , \tag{2.58}$$

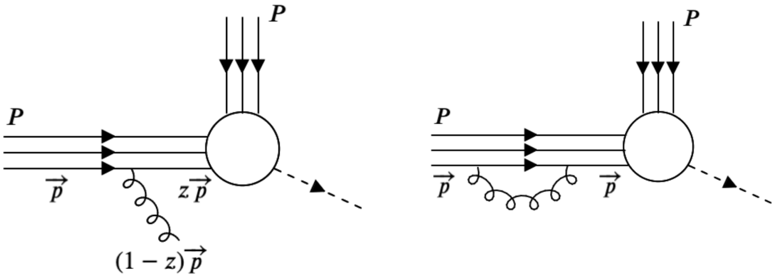


Figure 2.2.: Examples of Feynman diagrams for proton-proton collisions where a real (left) and virtual (right) gluon is emitted from the initial state. The cancellation between the two contributions holds only in the soft limit ($z \to 1$), as the momentum of the parton undergoing the hard scattering is unaltered. When we consider the collinear limit, the parton momentum changes ($\vec{p} \to z\vec{p}$) in the real contribution, while it remains the same in the virtual one ($\vec{p}$). This mismatch leads to a non-complete cancellation of collinear singularities.

where the subscript $i$ can take any possible flavour for both quarks and antiquarks. The DGLAP equation describes the evolution of the PDFs with the factorization scale: given a boundary condition at a scale $\mu_0$, equation (2.58) predicts the values of the PDFs at any other scale $\mu$. The $P_{ij}$ kernels are the *regularized Altarelli-Parisi splitting functions*, which are universal perturbative functions that admit an expansion in the strong coupling:

$$P_{ij}(z, \alpha_s(\mu^2)) = P_{ij}^{(0)}(z) + \frac{\alpha_s}{2\pi} P_{ij}^{(1)}(z) + \left(\frac{\alpha_s}{2\pi}\right)^2 P_{ij}^{(2)}(z) + \mathcal{O}(\alpha_s^3) \,. \tag{2.59}$$

They represent the probability that a parton $j$ emits a parton $i$ with energy fraction $z$. The LO expressions can be taken e.g. from [17]:

$$\begin{aligned} P_{q_i q_j}^{(0)}(z) &= \delta_{ij} C_\mathrm{F} \left( \frac{1+z^2}{[1-z]_+} + \frac{3}{2}\delta(1-z) \right), \\ P_{g q_i}^{(0)}(z) &= C_\mathrm{F} \left( \frac{1+(1-z)^2}{z} \right), \\ P_{q_i g}^{(0)}(z) &= \frac{1}{2}\left(z^2 + (1-z)^2\right), \\ P_{gg}^{(0)}(z) &= 2C_\mathrm{A} \left( \frac{z}{[1-z]_+} + (1-z)(z + \frac{1}{z}) \right) + \frac{11C_\mathrm{A} - 2n_f}{6}\,\delta(1-z)\,, \end{aligned} \tag{2.60}$$

where we used the *plus* distribution defined as:

$$\int_0^1 \mathrm{d}z \, \frac{f(z)}{[1-z]_+} = \int_0^1 \mathrm{d}z \, \frac{f(z) - f(1)}{1-z} \,. \tag{2.61}$$

This approach for handling collinear initial-state divergences is equivalent to the introduction of collinear counterterms at the level of the cross section. More precisely, reviewing the discussion presented in section 2.1 of [5], equation (2.14) now becomes:

$$\begin{aligned}\sigma_{\mathrm{NLO}} = \int \mathrm{d}\mathbf{\Phi}_n \mathcal{L}\Big[\mathcal{B}(\mathbf{\Phi}_n + \mathcal{V}_{\mathrm{b}}(\mathbf{\Phi}_n))\Big] &+ \int \mathrm{d}\mathbf{\Phi}_{n+1}\mathcal{L}R(\mathbf{\Phi}_{n+1}) \\ &+ \int \mathrm{d}\mathbf{\Phi}_{n,\oplus}\, \mathcal{L}\, \mathcal{G}_{\oplus,\mathrm{b}}(\mathbf{\Phi}_{n,\oplus}) + \int \mathrm{d}\mathbf{\Phi}_{n,\ominus}\, \mathcal{L}\, \mathcal{G}_{\ominus,\mathrm{b}}(\mathbf{\Phi}_{n,\ominus})\,, \end{aligned} \tag{2.62}$$

where the subscript b stands for *bare*, which means that IR divergences are still present. In the virtual correction, UV divergences have already been removed through renormalization. The two counterterms $\mathcal{G}_{\oplus,\mathrm{b}}$ and $\mathcal{G}_{\ominus,\mathrm{b}}$ are divergent in the collinear limit, therefore they are treated in dimensional regularization using $d = 4 + 2\epsilon$ for defining the dimensionality of the phase space. They thus present poles in $\epsilon$. The corresponding phase spaces $\mathbf{\Phi}_{n,\oplus}$ and $\mathbf{\Phi}_{n,\ominus}$ are given by configurations where one final-state parton is collinear to an incoming one. Defining with $z$ the fraction of momentum of the incoming particle after radiation, the two phase spaces can be written as:

$$\begin{aligned} \mathbf{\Phi}_{n,\oplus} &= \{x_\oplus, x_\ominus, z, k_1, ..., k_n\}, \qquad & z x_\oplus K_\oplus + x_\ominus K_\ominus &= \sum_i^n k_i\,, \\ \mathbf{\Phi}_{n,\ominus} &= \{x_\oplus, x_\ominus, z, k_1, ..., k_n\}, \qquad & x_\oplus K_\oplus + z x_\ominus K_\ominus &= \sum_i^n k_i\,. \end{aligned} \tag{2.63}$$

They are formally $n$-body configurations: we can introduce an underlying $n$-body phase space $\bar{\mathbf{\Phi}}_n$ as:

$$\begin{aligned} \mathbf{\Phi}_{n,\oplus}: \quad & \bar{\mathbf{\Phi}}_n = \{\bar{x}_\oplus, \bar{x}_\ominus, k_1, ..., k_n\} \quad \text{with} \quad \bar{x}_\oplus = z x_\oplus,\ \bar{x}_\ominus = x_\ominus\,, \\ \mathbf{\Phi}_{n,\ominus}: \quad & \bar{\mathbf{\Phi}}_n = \{\bar{x}_\oplus, \bar{x}_\ominus, k_1, ..., k_n\} \quad \text{with} \quad \bar{x}_\oplus = x_\oplus,\ \bar{x}_\ominus = z x_\ominus\,. \end{aligned} \tag{2.64}$$

Using this formalism, the expectation value of a generic IR-safe observable $\mathcal{O}$ can be obtained at NLO starting from (2.62) as:

$$\begin{aligned} \langle \mathcal{O} \rangle = \int \mathrm{d}\mathbf{\Phi}_n \mathcal{L}\, \mathcal{O}_n(\mathbf{\Phi}_n) \Big[\mathcal{B}(\mathbf{\Phi}_n) + \mathcal{V}_{\mathrm{b}}(\mathbf{\Phi}_n))\Big] & \\ + \int \mathrm{d}\mathbf{\Phi}_{n+1}\mathcal{L}\, \mathcal{O}_{n+1}(\mathbf{\Phi}_{n+1}) R(\mathbf{\Phi}_{n+1}) + \int \mathrm{d}\mathbf{\Phi}_{n,\oplus}\, \mathcal{L}\, \mathcal{O}_n(\bar{\mathbf{\Phi}}_n)\, \mathcal{G}_{\oplus,\mathrm{b}}(\mathbf{\Phi}_{n,\oplus}) & \\ + \int \mathrm{d}\mathbf{\Phi}_{n,\ominus} \mathcal{L}\, \mathcal{O}_n(\bar{\mathbf{\Phi}}_n)\, \mathcal{G}_{\oplus,\mathrm{b}}(\mathbf{\Phi}_{n,\ominus})\,. & \end{aligned} \tag{2.65}$$

Thanks to the IR-safety of $\mathcal{O}$, in the last two terms the argument of the observable is set equal to $\bar{\mathbf{\Phi}}_n$. If we require that the Born contribution $\mathcal{B}$ is finite, all but this term on the right-hand side of equation (2.65) are separately divergent, while the final result is finite. Equation (2.65) is thus not usable in a computer program, as it requires the separate calculation of divergent integrals. In order to overcome this problem, it is necessary to render separately finite all the contributions entering (2.65). In this context, one of the most successful procedures is given by a *subtraction method*.

Before discussing how a subtraction method is constructed, we make a final comment on the presence of the factorization scale in equation (2.4). On the same footing as the renormalization scale, the physical cross section does not depend on the factorization scale but, once the perturbative series is truncated at a certain perturbative order, a residual dependence on $\mu_F$ survives. In order to give an estimate of missing higher-order corrections, it is possible to set both the renormalization and the factorization scales to a typical energy scale $Q$ of the process

$$\mu_{R,0} = \mu_{F,0} \sim Q\,, \tag{2.66}$$

and vary them by an arbitrary factor around the central value:

$$\mu_R = K_R \cdot \mu_{R,0} \quad \text{and} \quad \mu_F = K_F \cdot \mu_{F,0}\,. \tag{2.67}$$

Customary, the $(K_R,\, K_F)$ combinations that are used are the following

$$(K_R,\, K_F) = \{(1,\, 1),\, (1,\, 1/2),\, (1/2,\, 1),\, (1,\, 2),\, (2,\, 1),\, (1/2,\, 1/2),\, (2,\, 2)\}\,, \tag{2.68}$$

in which we consider a factor of two above and below the central values, eliminating the cases in which the two scales are shifted in opposite directions (namely, $(1/2,\, 2)$ and $(2,\, 1/2)$). The result obtained using the central scales $(\mu_{R,0}\,, \mu_{F,0})$ represents the final prediction and the corresponding error bar is given by the envelope of the different results obtained with the combinations in equation (2.68). More precisely, among the seven predictions given by (2.68), the error bar is obtained by taking the maximum and the minimum values. This procedure is referred to as *7-point scale variation* and the physical meaning is the following: the final result should not manifest a strong dependence on $\mu_R$ and $\mu_F$, thus we expect that varying them arbitrarily does not change significantly the final prediction and higher-order corrections should be covered by the estimated the error bars. We stress that using large variation factors could lead to instabilities in the convergence of the series because of the appearance of large logarithmic contributions, and thus the value of 2 is commonly accepted.

### 2.1.4. Subtraction method

The expectation value of an IR-safe observable $\mathcal{O}$ at NLO accuracy can be obtained using equation (2.65). In general, the different ingredients in the formula are too complicated to be calculated analytically, therefore a numerical evaluation is needed. A numerical calculation is however not possible at the moment because IR divergences are still present in different contributions (virtual, real and counterterms) that are integrated over different phase spaces. In order to overcome this problem, we need to make explicitly finite all the terms appearing in our master formula (2.65). In this section, we review how to do that using a subtraction method. Notice that there are different subtraction schemes available in the literature and a detailed analysis of them is beyond the scope of this thesis. We thus review the general formalism of a local subtraction scheme, as it is the one implemented within the POWHEG framework [18], which will be used throughout this thesis. This section is inspired by section 2.2 of [5].

A local subtraction scheme consists in removing IR divergences point by point in the phase space summing and subtracting the same term to the virtual and real corrections so that the two contributions are separately finite. This implies that the chosen term must match the

IR singularities of the real and virtual corrections and must be analytically integrable over the radiation phase space. More precisely, the real contribution $R$ is associated to singular regions $\alpha$ in the IR limit which can refer to a:

- soft region (S), where the momentum of one of the final state partons goes to zero;
- final-state collinear region (FSC), where two massless final-state partons have parallel momenta;
- initial-state collinear region (ISC), where the momentum of one massless final-state parton is parallel to the momentum of one incoming parton.

A local subtraction method consists in identifying a real counterterm $\mathcal{C}^{(\alpha)}$ that shows the same divergent behaviour of the real contribution $R$ in the corresponding singular $\alpha$-region. To do so, we first identify a mapping $M^{(\alpha)}$ that maps a point of the $\mathbf{\Phi}_{n+1}$ phase space into a singular configuration $\alpha$, associated to the $\tilde{\mathbf{\Phi}}^{\alpha}_{n+1}$ phase space:

$$\tilde{\mathbf{\Phi}}^{(\alpha)}_{n+1} = M^{(\alpha)}(\mathbf{\Phi}_{n+1}), \quad \tilde{\mathbf{\Phi}}^{(\alpha)}_{n+1} = \{\tilde{x}^{(\alpha)}_{\oplus}, \tilde{x}^{(\alpha)}_{\ominus}, \tilde{k}^{(\alpha)}_{1}, ..., \tilde{k}^{(\alpha)}_{n+1}\}\,. \tag{2.69}$$

Each mapping $M^{(\alpha)}$ depends on the singular region $\alpha$ and the only requirement is that it must be smooth when approaching the singular region, where it becomes the identity. Divergences in the real contribution are thus cancelled as follows

$$R(\mathbf{\Phi}_{n+1})\mathcal{O}_{n+1}(\mathbf{\Phi}_{n+1}) - \sum_{\alpha}\mathcal{C}^{(\alpha)}(\mathbf{\Phi}_{n+1})\mathcal{O}_{n+1}(M^{(\alpha)}(\mathbf{\Phi}_{n+1}))\,, \tag{2.70}$$

with $\mathcal{O}$ a generic IR-safe observable.

Any singular configuration $\tilde{\mathbf{\Phi}}^{(\alpha)}_{n+1}$ can be associated to a $n$-body one $\bar{\mathbf{\Phi}}^{(\alpha)}_{n}$:

- $\alpha \in$ S : $\bar{\mathbf{\Phi}}^{(\alpha)}_{n}$ is equal to $\tilde{\mathbf{\Phi}}^{(\alpha)}_{n+1}$ removing the parton with momentum becoming soft;
- $\alpha \in$ FSC: $\bar{\mathbf{\Phi}}^{(\alpha)}_{n}$ is equal to $\tilde{\mathbf{\Phi}}^{(\alpha)}_{n+1}$ where the two collinear partons $i$ and $j$ are recombined together into a new parton $k$ so that $p_k = p_i + p_j$;
- $\alpha \in$ ISC: $\bar{\mathbf{\Phi}}^{(\alpha)}_{n}$ is equal to $\tilde{\mathbf{\Phi}}^{(\alpha)}_{n+1}$ where the collinear parton is removed and the momentum of the initial-state emitter is replaced with its momentum after radiation.

In all the three above cases, the final state partons are relabelled $\{1, ..., n+1\} \to \{1, ..., n\}$ and momentum conservation applies as:

$$\bar{x}_{\oplus}K_{\oplus} + \bar{x}_{\ominus}K_{\ominus} = \sum_{j=1}^{n}\bar{k}_j\,. \tag{2.71}$$

For S and FSC regions, the incoming momenta remain unaltered:

$$\bar{x}_{\oplus} = \tilde{x}_{\oplus}\,, \quad \bar{x}_{\ominus} = \tilde{x}_{\ominus}\,. \tag{2.72}$$

For ISC regions, one of the two incoming momenta $\bar{x}$ is smaller than the corresponding $\tilde{x}$:

$$\begin{aligned} &\oplus\ \text{case:} \qquad \bar{x}_{\oplus} < \tilde{x}_{\oplus}, \quad \bar{x}_{\ominus} = \tilde{x}_{\ominus}\,, \\ &\ominus\ \text{case:} \qquad \bar{x}_{\oplus} = \tilde{x}_{\oplus}, \quad \bar{x}_{\ominus} < \tilde{x}_{\ominus}\,. \end{aligned} \tag{2.73}$$

We can now perform explicitly the subtraction rewriting the integral of the real contribution $R$ in equation (2.65) as

$$\begin{aligned} \int \mathrm{d}\mathbf{\Phi}_{n+1} \mathcal{L}\, \mathcal{O}_{n+1}(\mathbf{\Phi}_{n+1})\, R(\mathbf{\Phi}_{n+1}) &= \sum_{\alpha} \int \mathrm{d}\mathbf{\Phi}_{n+1}\, \tilde{\mathcal{L}}^{(\alpha)}\, \mathcal{O}_n(\bar{\mathbf{\Phi}}_n^{(\alpha)})\, \mathcal{C}^{(\alpha)}(\mathbf{\Phi}_{n+1}) \\ &+ \int \mathrm{d}\mathbf{\Phi}_{n+1} \Big\{ \mathcal{L}\, \mathcal{O}_{n+1}(\mathbf{\Phi}_{n+1})\, R(\mathbf{\Phi}_{n+1}) - \sum_{\alpha} \tilde{\mathcal{L}}^{(\alpha)}\, \mathcal{O}_n(\bar{\mathbf{\Phi}}_n^{(\alpha)})\, \mathcal{C}^{(\alpha)}(\mathbf{\Phi}_{n+1}) \Big\}\,, \end{aligned} \tag{2.74}$$

where we introduced $\tilde{\mathcal{L}} = \mathcal{L}(\tilde{x}_{\oplus}, \tilde{x}_{\ominus})$ and we used the IR-safety of $\mathcal{O}$. The second line of equation (2.74) is finite and can be integrated in 4 dimensions over the $\mathbf{\Phi}_{n+1}$ phase space.

The first term on the r.h.s. of equation (2.74) is still divergent and, to handle it, we express the $(n+1)$-body phase space for each $\alpha$ as

$$\mathbf{\Phi}_{n+1} \overset{(\alpha)}{\longleftrightarrow} \Big\{ \bar{\mathbf{\Phi}}_n^{(\alpha)}, \Phi_{\text{rad}}^{(\alpha)} \Big\}, \qquad \mathrm{d}\mathbf{\Phi}_{n+1} = \mathrm{d}\bar{\mathbf{\Phi}}_n^{(\alpha)}\, \mathrm{d}\Phi_{\text{rad}}^{(\alpha)}\,, \tag{2.75}$$

which means that we describe the $(n+1)$-body phase space through the $n$-body phase space, constructed as described above, and a set of extra radiation variables that define the emission. At this point, we make the following considerations:

- When considering FSC and S regions, the luminosity in the $(n+1)$-body configuration is equal to the luminosity in the $n$-body configuration $\tilde{\mathcal{L}} = \mathcal{L}(\tilde{x}_{\oplus}, \tilde{x}_{\ominus}) = \bar{\mathcal{L}}(\bar{x}_{\oplus}, \bar{x}_{\ominus})$. Introducing the notation

  $$\bar{\mathcal{C}}^{(\alpha)}(\bar{\mathbf{\Phi}}_n^{(\alpha)}) = \int \mathrm{d}\Phi_{\text{rad}}^{(\alpha)}\, \mathcal{C}^{(\alpha)}(\mathbf{\Phi}_{n+1})\,, \tag{2.76}$$

  the divergent term in (2.74) becomes:

  $$\int \mathrm{d}\mathbf{\Phi}_{n+1} \bar{\mathcal{L}}^{(\alpha)}\, \mathcal{O}_n(\bar{\mathbf{\Phi}}_n^{(\alpha)})\, \mathcal{C}^{(\alpha)}(\mathbf{\Phi}_{n+1}) = \int \mathrm{d}\bar{\mathbf{\Phi}}_n^{(\alpha)}\, \tilde{\mathcal{L}}^{(\alpha)}\, \mathcal{O}_n(\bar{\mathbf{\Phi}}_n^{(\alpha)})\, \bar{\mathcal{C}}^{(\alpha)}(\bar{\mathbf{\Phi}}_n^{(\alpha)})\,. \tag{2.77}$$

- In the case of ISC regions, the luminosity $\tilde{\mathcal{L}}$ does not coincide with $\bar{\mathcal{L}}$. We thus define

  $$\begin{aligned} &\oplus\ \text{case:} \quad \bar{\mathcal{C}}^{(\alpha)}(\bar{\mathbf{\Phi}}_n^{(\alpha)}, z) = \int \mathrm{d}\Phi_{\text{rad}}^{(\alpha)}\, \mathcal{C}^{(\alpha)}(\mathbf{\Phi}_{n+1})\, z\, \delta(z - \bar{x}_{\oplus}/\tilde{x}_{\oplus})\,, \\ &\ominus\ \text{case:} \quad \bar{\mathcal{C}}^{(\alpha)}(\bar{\mathbf{\Phi}}_n^{(\alpha)}, z) = \int \mathrm{d}\Phi_{\text{rad}}^{(\alpha)}\, \mathcal{C}^{(\alpha)}(\mathbf{\Phi}_{n+1})\, z\, \delta(z - \bar{x}_{\ominus}/\tilde{x}_{\ominus})\,, \end{aligned} \tag{2.78}$$

  where $z$ is the momentum fraction of the incoming parton after radiation. The divergent term in (2.74) can now be rewritten as:

  $$\int \mathrm{d}\mathbf{\Phi}_{n+1} \bar{\mathcal{L}}^{(\alpha)}\, \mathcal{O}_n(\bar{\mathbf{\Phi}}_n^{(\alpha)})\, \mathcal{C}^{(\alpha)}(\mathbf{\Phi}_{n+1}) = \int \mathrm{d}\bar{\mathbf{\Phi}}_n^{(\alpha)} \frac{\mathrm{d}z}{z}\, \tilde{\mathcal{L}}^{(\alpha)}\, \mathcal{O}_n(\bar{\mathbf{\Phi}}_n^{(\alpha)})\, \bar{\mathcal{C}}^{(\alpha)}(\bar{\Phi}_n^{(\alpha)})\,. \tag{2.79}$$

Note that the following relations hold:

$$
\begin{aligned}
&\oplus \text{ case:} \qquad && \mathrm{d}\mathbf{\Phi}_{n,\oplus} = \mathrm{d}\bar{\mathbf{\Phi}}_n \frac{\mathrm{d}z}{z}\,, \\
&\ominus \text{ case:} \qquad && \mathrm{d}\mathbf{\Phi}_{n,\ominus} = \mathrm{d}\bar{\mathbf{\Phi}}_n \frac{\mathrm{d}z}{z}\,. \qquad (2.80)
\end{aligned}
$$

The full NLO master formula (2.65) now becomes:

$$
\begin{aligned}
\langle \mathcal{O} \rangle = & \int \mathrm{d}\mathbf{\Phi}_n \, \mathcal{L} \, \mathcal{O}_n(\mathbf{\Phi}_n) \Big[ \mathcal{B}(\mathbf{\Phi}_n) + \mathcal{V}_{\mathrm{b}}(\mathbf{\Phi}_n) ) \Big] \\
& + \int \mathrm{d}\mathbf{\Phi}_{n+1} \Big\{ \mathcal{L} \, \mathcal{O}_{n+1}(\mathbf{\Phi}_{n+1}) R(\mathbf{\Phi}_{n+1}) - \sum_{\alpha} \big[ \tilde{\mathcal{L}}^{(\alpha)} \, \mathcal{O}_n(\bar{\mathbf{\Phi}}_n^{(\alpha)}) \mathcal{C}^{(\alpha)}(\mathbf{\Phi}_{n+1}) \big] \Big\} \\
& + \sum_{\alpha \in \{\mathrm{S,FSC}\}} \int \mathrm{d}\bar{\mathbf{\Phi}}_n^{(\alpha)} \, \tilde{\mathcal{L}}^{(\alpha)} \mathcal{O}_n(\bar{\mathbf{\Phi}}_n^{(\alpha)}) \, \bar{\mathcal{C}}^{(\alpha)}(\bar{\mathbf{\Phi}}_n^{(\alpha}) \\
& + \sum_{\alpha \in \{\mathrm{ISC}_{\oplus/\ominus}\}} \int \mathrm{d}\bar{\mathbf{\Phi}}_{n,\oplus/\ominus}^{(\alpha)} \, \tilde{\mathcal{L}}^{(\alpha)} \, \mathcal{O}_n(\bar{\mathbf{\Phi}}_n^{(\alpha)}) \, \bar{\mathcal{C}}^{(\alpha)}(\bar{\mathbf{\Phi}}_{n,\oplus/\ominus}^{(\alpha)}) \\
& + \int \mathrm{d}\mathbf{\Phi}_{n,\oplus} \, \tilde{\mathcal{L}} \, \mathcal{O}_n(\bar{\mathbf{\Phi}}_n) \, \mathcal{G}_{\oplus,\mathrm{b}}(\mathbf{\Phi}_{n,\oplus}) + \int \mathrm{d}\mathbf{\Phi}_{n,\ominus} \, \tilde{\mathcal{L}} \, \mathcal{O}_n(\bar{\mathbf{\Phi}}_n) \, \mathcal{G}_{\ominus,\mathrm{b}}(\mathbf{\Phi}_{n,\ominus})\,. \qquad (2.81)
\end{aligned}
$$

In the last line we have substituted $\mathcal{L}$ with $\tilde{\mathcal{L}}$: this is correct since $x_{\oplus/\ominus} = \tilde{x}_{\oplus/\ominus}$ in the phase space of the counterterms, as the mapping (2.69) is the identity in the singular region. For ISC regions, it is always possible to write

$$
\mathcal{G}_{\oplus,\mathrm{b}}(\mathbf{\Phi}_{n,\oplus}) + \sum_{\alpha \in \{\mathrm{ISC}_\oplus\}} \bar{C}^{(\alpha)}(\mathbf{\Phi}_{n,\oplus}) = \mathcal{G}_{\oplus}(\mathbf{\Phi}_{n,\oplus}) + \delta(1-z) \, \mathcal{G}_{\oplus}^{\mathrm{div}}(\bar{\mathbf{\Phi}}_n)\,, \qquad (2.82)
$$

$$
\mathcal{G}_{\ominus,\mathrm{b}}(\mathbf{\Phi}_{n,\ominus}) + \sum_{\alpha \in \{\mathrm{ISC}_\ominus\}} \bar{C}^{(\alpha)}(\mathbf{\Phi}_{n,\ominus}) = \mathcal{G}_{\ominus}(\mathbf{\Phi}_{n,\ominus}) + \delta(1-z) \, \mathcal{G}_{\ominus}^{\mathrm{div}}(\bar{\mathbf{\Phi}}_n)\,, \qquad (2.83)
$$

in which we explicitly separate finite contributions ($\mathcal{G}_{\oplus}(\mathbf{\Phi}_{n,\oplus})$ and $\mathcal{G}_{\ominus}(\mathbf{\Phi}_{n,\ominus})$) from terms that still contain poles in $\epsilon$ ($\mathcal{G}_{\oplus}^{\mathrm{div}}(\bar{\Phi}_n)$ and $\mathcal{G}_{\ominus}^{\mathrm{div}}(\bar{\mathbf{\Phi}}_n)$). These poles have a soft origin. Moreover, the expression

$$
\mathcal{V}(\mathbf{\Phi}_n) = \mathcal{V}_{\mathrm{b}}(\mathbf{\Phi}_n) + \sum_{\alpha \in \{\mathrm{FSR,S}\}} \Big[ \bar{\mathcal{C}}^{(\alpha)}(\bar{\mathbf{\Phi}}_n) + \mathcal{G}_{\oplus}^{\mathrm{div}}(\bar{\mathbf{\Phi}}_n) + \mathcal{G}_{\ominus}^{\mathrm{div}}(\bar{\mathbf{\Phi}}_n) \Big]^{\bar{\mathbf{\Phi}}_n = \mathbf{\Phi}_n} \qquad (2.84)
$$

is finite. The notation

$$
[...]^{\bar{\mathbf{\Phi}}_n = \mathbf{\Phi}_n}
$$

means that all the terms in brackets are evaluated for phase space variables $\bar{\mathbf{\Phi}}_n$ equal to $\mathbf{\Phi}_n$.

For ease of notation, we redefine all the NLO ingredients reabsorbing the luminosity factors as

$$
B \equiv \mathcal{L}\,\mathcal{B}, \;\; R \equiv \mathcal{L} R, \;\; V \equiv \mathcal{L}\,\mathcal{V}, \;\; C^{(\alpha)} \equiv \tilde{\mathcal{L}}^{(\alpha)}\, C^{(\alpha)}, \;\; \mathcal{G}_{\oplus} = \tilde{\mathcal{L}}\mathcal{G}_{\oplus}, \;\; \mathcal{G}_{\ominus} = \tilde{\mathcal{L}}\mathcal{G}_{\ominus}\,,
$$

so that equation (2.81) takes its final form:

$$\begin{aligned}\langle \mathcal{O} \rangle = &\int \mathrm{d}\mathbf{\Phi}_n \mathcal{O}_n(\mathbf{\Phi}_n) \Big[ B(\mathbf{\Phi}_n) + V(\mathbf{\Phi}_n) \Big] \\ &+ \int \mathrm{d}\mathbf{\Phi}_{n+1} \Big\{ \mathcal{O}_{n+1}(\mathbf{\Phi}_{n+1}) R(\mathbf{\Phi}_{n+1}) - \sum_\alpha \left[ \mathcal{O}_n(\bar{\mathbf{\Phi}}_n^\alpha) \mathcal{C}^{(\alpha)}(\mathbf{\Phi}_{n+1}) \right] \Big\} \\ &+ \int \mathrm{d}\mathbf{\Phi}_{n,\oplus} \mathcal{O}_n(\bar{\mathbf{\Phi}}_n) \mathcal{G}_\oplus(\mathbf{\Phi}_{n,\oplus}) + \int \mathrm{d}\mathbf{\Phi}_{n,\ominus} \mathcal{O}_n(\bar{\mathbf{\Phi}}_n) \mathcal{G}_\ominus(\mathbf{\Phi}_{n,\ominus}) \,. \end{aligned} \tag{2.85}$$

All the terms appearing in the previous equation are finite in 4 dimensions and can thus be integrated numerically.

Different methods are available in the literature for performing the subtraction of IR singularities at NLO accuracy [19, 20, 21, 22] and they differ for the specific choice of the counterterms $C^{(\alpha)}$. Anyway, they share the same structure discussed above. In the POWHEG framework the subtraction is performed through the FKS method [21, 22] in which the real contribution is split into different singular regions where at most one parton becomes soft and/or collinear. Each region is then treated separately. We highlight that nowadays the local subtraction of IR poles at NLO accuracy is fully understood and completely automatized, while an extension to NNLO accuracy is still highly non-trivial. At this accuracy, many methods have been proposed in the last years [23, 24, 25, 26, 27, 28], but automation is still not possible.

## 2.2. Parton Showers

In the previous sections we discussed in detail how to use fixed-order calculations in perturbative QCD for obtaining predictions at colliders. Despite the enormous progress done in the last years, this approach is not always optimal and presents some intrinsic limitations. Firstly, this approach relies entirely on our ability in calculating Feynman diagrams: the more complex the process, the more involved the calculation. Moreover, the complete cancellation of divergences is fulfilled only when studying inclusive (IR-safe) observables. In a realistic description of a collider event, we may want to be more exclusive on the final state, e.g. restricting the phase space of the process or applying kinematic cuts in order to perform a comparison with data. Anytime we want to be more exclusive on the final state, the fixed-order approach is going to fail because of the appearance of large logarithmic contributions that spoil the perturbative convergence of the series. More precisely, considering exclusive observables introduces new scales $Q_i$ in the process that directly affect the definition of the observables under study. The integral of the real contribution thus depends directly on these new scales, while the virtual contribution remains unaffected, leading to a partial cancellation of divergences in the IR limit. What comes out is a logarithmic residue of the form $L = \log Q_i/Q$, where $Q$ is the hard scale of the process. Considering that IR singularities are of soft and/or collinear origins, we have at most two powers of this potentially large logarithms for a single power of the strong coupling $(\alpha_s L^2)$. To all orders, we thus expect terms of the form $(\alpha_s L^2)^n$ to show up. Anytime the scales are in the regime $\alpha_s \sim L^{-2}$, the perturbative series breaks down. In order to restore the predictivity of the calculation, we need to account for these enhanced contributions to all orders through *resummation*.

Resummation consists in reorganising the perturbative series in terms of large logarithmic contributions: the dominant contributions (e.g. $(\alpha_s L^2)^n$) are called *leading logarithmic* terms (LL), contributions with one less power of the logarithm (e.g. $(\alpha_s L)^n$) are *next-to-leading logarithmic* terms (NLL), and so on. When considering an appropriate range for $L$, the resummed hierarchy is convergent, which means that LL contributions are larger than NLL, which are larger than NNLL, etc. This resummation can be performed either analytically, like in [29, 30, 31], or numerically through a parton shower simulation implemented within a General-purpose Monte Carlo (GPMC) event generator, like PYTHIA [32], HERWIG [33] and SHERPA [34].

GPMC event generators are among the most employed tools in the high-energy physics community as they provide a fully exclusive picture of high-energy collisions in a flexible way. They describe the evolution of a collider event from hard scales to the detector level, accounting for all the different stages that constitute an event, both in the perturbative and non-perturbative regimes. GPMC event generators thus deal with both short-distance physics, where perturbation theory is allowed, and long-distance effects, where phenomenological models tuned to data are needed.

In a GPMC event generator, the first step is the generation of the hard scattering, considered at LO accuracy only, and then subsequent splittings in the collinear approximation are produced. The hard scattering (in a simple case it can be a $2 \to 2$ process) is thus embedded into a more complicated $2 \to n$ process, in which the additional partons are produced by both initial and final state radiations in the collinear approximation. These additional splittings determine a degradation of the energy and the process stops when the non-perturbative regime is approached at the QCD cutoff $\Lambda_{\text{QCD}}$. Once the evolution of the shower ends, the numerous final-state partons are converted into hadrons according to appropriate phenomenological models. The generated cascade of particles from the high-energy limit down to the non-perturbative scale is the parton shower. Through the parton shower, all classes of (potentially) large logarithmic contributions are cancelled, but the accuracy of this resummation procedure is low, only LL.

A parton shower algorithm is based on the general property of collinear factorization of QCD: in the collinear limit, QCD squared matrix elements factorize in a universal way as shown in equation (2.53). The universal probability of emitting a collinear parton can thus be used as a building block for constructing an iterative algorithm that generates the emission of $n$ extra collinear partons on top of the hard scattering. We now discuss the basic features of parton shower simulations as presented in [35], where the interested reader can find a complete pedagogical review of the topic.

Parton shower simulations deal separately with final-state radiation (called *time-like* shower) and initial-state radiation (called *space-like* shower) and we can start our discussion analysing the simpler case of final-state radiation. The probability of a parton $i$ to undergo a collinear splitting $i \to jk$ is given by (see equation (2.48) and the subsequent discussion)

$$\mathrm{d}P_i(z,t) = \frac{\alpha_s(t)}{2\pi}\frac{\mathrm{d}t}{t}\mathrm{d}z\hat{P}_{ji}(z)\,, \tag{2.86}$$

where $z$ is the energy fraction of parton $j$, $t$ is an appropriate scale of the process referred to as *hardness* and $\hat{P}_{ji}(z)$ is the unregularized Altarelli-Parisi splitting kernel defined in equation (2.55). The choice of the hardness $t$ is not unique: it is natural to associate to $t$ the virtuality of an intermediate off-shell parton $m^2$, but other options are possible. In fact, if $t = f(z)m^2$, for a nice choice of the function $f(z)$ we obtain $\mathrm{d}t/t\,\mathrm{d}z = \mathrm{d}m^2/m^2\,\mathrm{d}z$. Nowadays, three different options are widely used: in early versions of SHERPA and PYTHIA, $t$ is the virtuality of an intermediate parton; in current versions of SHERPA and PYTHIA it is the transverse momentum of the emission and in HERWIG it is the emitting angle.

Divergences are eliminated considering that a physical measurement cannot distinguish between a pair of collinear partons and a single parton with the same total momentum. We can thus introduce a resolution criterion that allows one to generate only distributions of resolvable partons. For example, a possible choice is given by transverse momentum: two partons are resolvable if the relative transverse momentum is above a certain threshold. Concretely, the cutoff on the transverse momentum defines a kinematic range for $z$ and we can thus simply integrate over the kinematically allowed phase space:

$$\mathrm{d}P_i(t) = \frac{\alpha_s(t)}{2\pi}\frac{\mathrm{d}t}{t}\int_{z_{min}}^{z_{max}} \mathrm{d}z\hat{P}_{ji}(z)\,. \tag{2.87}$$

In this way, we eliminate both soft and collinear divergences and we obtain a total probability that is finite. This procedure breaks unitarity, as we are not including in the computation unresolved emissions and virtual corrections. This problem can be simply overcome considering that the non-emission probability can be calculated as one minus the emission probability, thanks to unitarity. Taking into account one emission in the infinitesimal hardness range $\mathrm{d}t$, this is $1 - \mathrm{d}P_i(t)$. If we introduce an ordering in the hardness $t$, we can now obtain the probability of not emitting a parton in between two scales $t_{min}$ and $t_{max}$: this probability is called *Sudakov form factor*.

With multiple emissions, the Sudakov form factor exponentiates as

$$\Delta_i(t_{min}, t_{max}) = \exp\left( -\int_{t_{min}}^{t_{max}} \mathrm{d}P_i(t)\right), \tag{2.88}$$

which means that, if we start the evolution from a scale $t_{max}$, the differential probability that the parton $i$ branches at a scale equal to $t_i$ is given by $\mathrm{d}P_i(t_i)\Delta_i(t_i, t_{max})$. Once $i$ has branched, the two daughters can branch in their turn, using as starting scale $t_{max} = t_i$. Iterating this procedure, we can construct the entire parton shower. This description, which is typical of final-state radiation, is called *forward evolution*, as the hardness $t$ evolves according to physical time.

The probabilistic description associated to Sudakov form factors renders the implementation of the above procedure particularly suitable for MC calculations. In fact, we can generate a random number $\rho$ in [0,1] and solve $\Delta_i(t_{max}, t_i) = \rho$ for $t_i$: if $t_i$ is greater than the PS cutoff, a resolvable emission is generated, otherwise no splitting is produced and the evolution of the shower ends. If a splitting is generated, the procedure is iterated setting as new maximum scale $t_{max} = t_i$.

As for the case of initial-state radiation, the construction of a PS could in principle follow the procedure used for final-state radiation, with the difference that emissions are now associated to possible splittings of the incoming partons before entering the hard collision. We could thus generate $1 \to 2$ splittings multiple times and two of the final generated partons enter the hard process, while the others contribute only as extra radiation. This procedure is anyway extremely inefficient, as it is rare to generate the kinematics suitable to produce a hard scattering. For this reason, PS coming from initial state particles is associated to a *backward evolution*, in the sense that we start from the hard scattering and we reconstruct the possible emissions moving backwards in time.

The main difference with the forward evolution shower is the definition of the Sudakov form factor. The DGLAP equation defined in (2.58) represents the probability of a parton $i$ to undergo a splitting $i \to jk$ when the energy is varied. From the point of view of backward evolution, starting from the DGLAP equation we can construct the relative probability of the parton $i$ to become unresolved as follows

$$\mathrm{d}P_i(t) = \frac{\mathrm{d}f_i(x,t)}{f_i(x,t)} = \frac{\alpha_s(t)}{2\pi}\frac{|\mathrm{d}t|}{t}\int_{z_{min}}^{z_{max}} \frac{\mathrm{d}z}{z}\frac{f_j(z,t)}{f_i(x,t)}\hat{P}_{ji}\left(\frac{x}{z}\right), \tag{2.89}$$

so that the cumulative effect of different backward emissions are described by a Sudakov form factor of the form

$$\Delta_i(t_{min}, t_{max}) = \exp\left(-\int_{t_{min}}^{t_{max}} \mathrm{d}P_i(t)\right), \tag{2.90}$$

where $\mathrm{d}P_i(t)$ is now defined in (2.89).

Up to now, we discussed the construction of a PS starting from the collinear limit of QCD amplitudes, but it is clear that also soft enhancement should be taken into account. It is possible to show that a PS in which the hardness $t$ is defined as the emitting angle or the transverse momentum already exhibits the correct behaviour in the soft regions [36, 37].

With the ingredients discussed above, we can now obtain the expectation value of an observable $\mathcal{O}$ in a PS framework for a generic $2 \to n$ process as follows: first, we consider all Born diagrams and at each vertex we include an appropriate splitting factor:

$$\frac{\alpha_s(t)}{2\pi}\frac{\mathrm{d}t}{t}\mathrm{d}z\hat{P}_{ji}(z)\,. \tag{2.91}$$

For each internal parton line $i$, we include the probability of not emitting any parton in between two scales, $t_1$ and $t_2$, which corresponds to the Sudakov form factor $\Delta_i(t_1, t_2)$. In case we are considering a final line, $t_2$ is set to the PS cutoff $t_0$.

If we consider only the first emission of the shower, we can obtain the total cross section as

$$\sigma^{PS} = \int \mathrm{d}\mathbf{\Phi}_n B(\mathbf{\Phi}_n)\left\{\Delta(t,t_0) + \int_{t_0}^{t}\frac{\mathrm{d}t'}{t'}\mathrm{d}z\,\frac{\alpha_s(t')}{2\pi}\hat{P}(z)\,\Delta(t,t')\right\}, \tag{2.92}$$

and, correspondingly, the expectation value of an observable $\mathcal{O}$ is given by:

$$\langle\mathcal{O}\rangle^{PS} = \int \mathrm{d}\mathbf{\Phi}_n B(\mathbf{\Phi}_n)\left\{\mathcal{O}_n(\mathbf{\Phi}_n)\,\Delta(t,t_0) + \int_{t_0}^{t}\frac{\mathrm{d}t'}{t'}\mathrm{d}z\,\frac{\alpha_s(t')}{2\pi}\hat{P}(z)\,\Delta(t,t')\mathcal{O}_{n+1}(\mathbf{\Phi}_n, \Phi_{rad})\right\}. \tag{2.93}$$

Note that the flavour indices have been removed for the ease of notation. We stress that all flavour configurations are considered in the splitting kernels and the Sudakov form factors. The cross section is thus made up of two terms: the first one is the Born level, in which no emission has been generated by the shower; the second one is given by the Born level times the emission of the first radiation in a PS approximation. Thanks to the unitarity of the parton shower, the total normalization is not changed, which means that the shower does not affect the inclusive cross section, but it has an impact on the shape of distributions. For generating multiple splittings and constructing the entire shower, we can simply manipulate the terms in the curly brackets iterating the procedure in a similar way.

Parton shower simulations generate only an approximation of real emissions, and unresolved and virtual corrections are encoded in the computation employing unitarity requirements. By construction, it does not include higher-order corrections with full matrix elements, which are needed for precision phenomenology. A PS simulation is thus suitable to obtain a realistic description of an event at colliders, but the accuracy of the prediction is low (in general, only LL). In the next chapter we will investigate the problem of matching the two approaches described in this chapter, namely fixed-order calculations and parton showers, in order to construct a single framework that keeps the best features of both. The underlying idea is retaining the high accuracy of the fixed-order computation to describe the hard scattering, maintaining a realistic and flexible picture of the entire event as given by the parton shower.

# 3. Matching fixed-order calculations with parton showers

In the past few decades, matching fixed-order calculations with parton showers became a crucial ingredient for precise simulations of hard processes at colliders. In the last chapter, we reviewed the main features of these two tools, showing that they are complementary approaches that can be fully exploited when matched together. The matching procedure consists in describing the hard scattering through the correct matrix elements, so that we can retain the highest accuracy possible in the cross section, without spoiling the LL structure of the parton shower. This procedure is not trivial because real matrix elements generate emissions not only in the hard regions but also in the soft and collinear regimes, where they are already accounted for by the PS in an approximate form. This double-counting problem already arises at NLO accuracy and it clearly becomes more complicated when higher-order corrections are included.

The current state of the art for precise theoretical predictions at colliders is given by NNLO+PS accuracy, which means that we keep the NNLO description of the hard scattering without ruining the parton shower simulation. The aim of this thesis is to present how to target NNLO+PS calculations using the MiNNLO$_{\mathrm{PS}}$ method [8].

In order to understand how to reach this level of precision, we first review how this problem was solved in the simpler NLO case, where the matching is now fully understood and automatized. In particular, two main solutions were proposed: MC@NLO [38] and POWHEG [18]. Even if MC@NLO is nowadays widely applicable and used, in this chapter we will present the main features of the POWHEG approach only, as it is the general framework in which the MiNNLO$_{\mathrm{PS}}$ method is implemented.

Once the matching problem was solved at NLO accuracy, targeting NNLO+PS calculations was however not straightforward. POWHEG was improved in the so-called MiNLO$'$ method[1] [39, 40], which provides a method to retain NLO accuracy in different jet multiplicities, as shown in table 3.1. More precisely, we can consider the production of a generic colour singlet F and study the accuracy we obtain when getting more and more exclusive over QCD radiation. Using the POWHEG method, we obtain NLO accuracy in distributions that are inclusive in F, LO accuracy in the first emission FJ and only LL accuracy starting from the second emission. When adopting the MiNLO$'$ approach, we maintain NLO accuracy in both F and FJ at the same time. The last step for obtaining a NNLO+PS description of the process consists in pushing the accuracy in distributions inclusive in F towards NNLO accuracy, which is done by the MiNNLO$_{\mathrm{PS}}$ method [8].

---

[1] The POWHEG framework was firstly improved in the MiNLO approach, in which no claim on the accuracy on

| | **F** | **FJ** | **FJJ** | **F(≥3J)** |
|---|---|---|---|---|
| POWHEG | NLO | LO | LL | LL |
| MiNLO′ | NLO | NLO | LO | LL |
| MiNNLO$_{\mathrm{PS}}$ | NNLO | NLO | LO | LL |

Table 3.1.: Formal accuracy reached through POWHEG, MiNLO′ and MiNNLO$_{\mathrm{PS}}$ for the production of a colour singlet F in association with extra QCD emissions J.

This chapter is structured as follows: we first present a review of the POWHEG method in section 3.1 and of the MiNLO and MiNLO′ procedures in section 3.2. In these sections, we highlight the main features of the different methods and we thus refer the interested reader to the appropriate references for more details. We then present a detailed overview of the MiNNLO$_{\mathrm{PS}}$ approach in section 3.3.

## 3.1. The POWHEG method

The POWHEG method aims to match fixed-order calculations with parton showers at NLO accuracy. The main issue is avoiding double counting, as the real emission generated by POWHEG with the correct matrix element is already accounted for in an approximate way by the shower.

The name POWHEG stands for *POsitive Weight Hardest Event Generator*, which is inspired by the procedure used for solving the double counting issue: POWHEG generates the first emission, which is the hardest, with the correct matrix element and then the parton shower can generate only softer extra radiation. The method concretely requires the application of a $p_{\mathrm{T}}$-veto: the transverse momentum of all the emissions produced by the shower (${p_{\mathrm{T}}}^{PS}$) must be smaller than the transverse momentum of the emission generated by POWHEG, commonly called `scalup`, at parton level (from now on, *Les Houches Event* (LHE) level):

$$ {p_{\mathrm{T}}}^{PS} < \texttt{scalup}\,. \tag{3.1}$$

This procedure is straightforward with $p_{\mathrm{T}}$-ordered showers, while angular-ordered showers need the inclusion of truncated showers for restoring soft coherence [18].

We now sketch how to explicitly construct the POWHEG master formula for the calculation of observables at NLO+PS accuracy following the discussion presented in [18, 5, 41]. We consider the production of a generic color singlet F and we define its Born phase space as $\mathbf{\Phi}_{\mathrm{F}}$. With J we indicate the extra emission associated to real radiation. Given an IR-safe observable $\mathcal{O}$, we can calculate its expectation value at NLO as follows

$$\langle \mathcal{O} \rangle = \int \mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\, \mathcal{O}(\mathbf{\Phi}_{\mathrm{F}})\ [B(\mathbf{\Phi}_{\mathrm{F}}) + V_{\mathrm{b}}(\mathbf{\Phi}_{\mathrm{F}})] + \int \mathrm{d}\mathbf{\Phi}_{\mathrm{FJ}}\, \mathcal{O}(\mathbf{\Phi}_{\mathrm{FJ}})\, R(\mathbf{\Phi}_{\mathrm{FJ}})\,, \tag{3.2}$$

---

the 0-jet bin was made.

where the subscript b in the virtual term means that it is still IR divergent (but UV finite). IR divergences are regularized through a subtraction scheme[2], so that we define an appropriate set of counterterms that separately eliminates IR divergences in the real and virtual contributions:

$$\begin{aligned}\langle\mathcal{O}\rangle = \int \mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\, \mathcal{O}(\mathbf{\Phi}_{\mathrm{F}}) \left[B(\mathbf{\Phi}_{\mathrm{F}}) + V_{\mathrm{b}}(\mathbf{\Phi}_{\mathrm{F}}) + \int \mathrm{d}\mathbf{\Phi}_{rad}\, C(\mathbf{\Phi}_{\mathrm{F}}, \Phi_{rad})\right] \\ + \int \mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\, \mathrm{d}\Phi_{rad} \left[\mathcal{O}(\mathbf{\Phi}_{\mathrm{F}}, \Phi_{rad})\, R(\mathbf{\Phi}_{\mathrm{FJ}}) - \mathcal{O}(\mathbf{\Phi}_{\mathrm{F}}, \Phi_{rad})\, C(\mathbf{\Phi}_{\mathrm{F}}, \Phi_{rad})\right] . \quad (3.3)\end{aligned}$$

Note that we do not include collinear counterterms (terms $\mathcal{G}_{\oplus/\ominus}$ in (2.81)) associated to initial-state emissions for ease of notation, but they are understood throughout the discussion. Moreover, we introduced a mapping between the F and the FJ phase spaces

$$\mathbf{\Phi}_{\mathrm{FJ}} = \mathbf{\Phi}_{\mathrm{F}}\, \Phi_{rad}\,, \tag{3.4}$$

so that the FJ phase space can be written in terms of the F phase space and a radiation phase space. Reabsorbing IR divergences into a redefinition of the virtual contribution

$$V(\mathbf{\Phi}_{\mathrm{F}}) = V_{\mathrm{b}}(\mathbf{\Phi}_{\mathrm{F}}) + \int \mathrm{d}\Phi_{rad}\, C(\mathbf{\Phi}_{\mathrm{F}}, \Phi_{rad})\,, \tag{3.5}$$

we obtain the following NLO master formula

$$\begin{aligned}\langle\mathcal{O}\rangle = \int \mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\, \mathcal{O}(\Phi_{\mathrm{F}}) \left[B(\mathbf{\Phi}_{\mathrm{F}}) + V(\mathbf{\Phi}_{\mathrm{F}})\right] \\ + \int \mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\, \mathrm{d}\Phi_{rad} \left[\mathcal{O}(\mathbf{\Phi}_{\mathrm{F}}, \Phi_{rad})\, R(\mathbf{\Phi}_{\mathrm{FJ}}) - \mathcal{O}(\mathbf{\Phi}_{\mathrm{F}}, \Phi_{rad})\, C(\mathbf{\Phi}_{\mathrm{F}}, \Phi_{rad})\right] , \quad (3.6)\end{aligned}$$

in which all the terms are finite. We recall that in the soft and/or collinear regions, the following equation is valid

$$\mathcal{O}(\mathbf{\Phi}_{\mathrm{F}}, \Phi_{rad}) = \mathcal{O}(\mathbf{\Phi}_{\mathrm{F}})\,, \tag{3.7}$$

thanks to the IR-safety of $\mathcal{O}$.

The key ingredient of the POWHEG method, namely $\bar{B}$, is obtained as

$$\bar{B}(\mathbf{\Phi}_{\mathrm{F}}) = B(\mathbf{\Phi}_{\mathrm{F}}) + V(\mathbf{\Phi}_{\mathrm{F}}) + \int \mathrm{d}\Phi_{rad} \left[R(\mathbf{\Phi}_{\mathrm{FJ}}) - C(\mathbf{\Phi}_{\mathrm{F}}, \Phi_{rad})\right] , \tag{3.8}$$

so we can thus rewrite equation (3.6) in the following form

$$\begin{aligned}\langle\mathcal{O}\rangle &= \int \mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\, \mathcal{O}(\mathbf{\Phi}_{\mathrm{F}}) \left[B(\mathbf{\Phi}_{\mathrm{F}}) + V(\mathbf{\Phi}_{\mathrm{F}}) + \int \mathrm{d}\Phi_{rad}\big(R(\mathbf{\Phi}_{\mathrm{FJ}}) - C(\mathbf{\Phi}_{\mathrm{F}}, \Phi_{rad})\big)\right] \\ &\quad + \int \mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\, \mathrm{d}\Phi_{rad}\, R(\mathbf{\Phi}_{\mathrm{FJ}}) \left[\mathcal{O}(\mathbf{\Phi}_{\mathrm{F}}, \Phi_{rad}) - \mathcal{O}(\mathbf{\Phi}_{\mathrm{F}})\right] \\ &= \int \mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\, \mathcal{O}(\mathbf{\Phi}_{\mathrm{F}}) \bar{B}(\mathbf{\Phi}_{\mathrm{F}}) + \int \mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\, \mathrm{d}\Phi_{rad}\, R(\mathbf{\Phi}_{\mathrm{FJ}}) [O(\mathbf{\Phi}_{\mathrm{F}}, \Phi_{rad}) - O(\mathbf{\Phi}_{\mathrm{F}})] . \end{aligned} \tag{3.9}$$

---

[2] In the POWHEG framework, the FKS method [21, 22] is employed.

This equation is written in a particularly convinient way for making a connection with the same expression in a PS algorithm. In fact, in a PS approach the expectation value of an IR-safe observable $\mathcal{O}$ is

$$\langle\mathcal{O}\rangle^{PS} = \int \mathrm{d}\mathbf{\Phi}_\mathrm{F} B(\mathbf{\Phi}_\mathrm{F}) \left\{ \Delta(Q,Q_0)\, \mathcal{O}(\mathbf{\Phi}_\mathrm{F}) + \int_{Q_0}^{Q} \mathrm{d}\Phi_{rad}\, \mathcal{K}(\Phi_{rad})\, \Delta(Q,t)\, \mathcal{O}(\mathbf{\Phi}_\mathrm{F},\Phi_{rad}) \right\}, \tag{3.10}$$

where we considered only the first emission (see equation (2.93)). $\Delta$ is the usual Sudakov form factor and the factor $\mathcal{K}$ contains the splitting kernels. The scales $Q$ and $Q_0$ define respectively the hard and the cutoff scales. We can expand this equation in terms of the strong coupling $\alpha_s$ up to NLO accuracy:

$$\begin{aligned}\langle\mathcal{O}\rangle^{\mathrm{PS}} &= \int \mathrm{d}\mathbf{\Phi}_\mathrm{F}\, B(\mathbf{\Phi}_\mathrm{F}) \left\{ \mathcal{O}(\mathbf{\Phi}_\mathrm{F}) + \int_{Q_0}^{Q} \mathrm{d}\Phi_{rad} \mathcal{K}(\Phi_{rad}) \left[\mathcal{O}(\mathbf{\Phi}_\mathrm{F},\Phi_{rad}) - \mathcal{O}(\mathbf{\Phi}_\mathrm{F})\right] \right\} \\ &= \int \mathrm{d}\mathbf{\Phi}_\mathrm{F}\, \mathcal{O}(\mathbf{\Phi}_\mathrm{F}) B(\mathbf{\Phi}_\mathrm{F}) + \int \mathrm{d}\mathbf{\Phi}_\mathrm{F}\, \mathrm{d}\Phi_{rad} B(\mathbf{\Phi}_\mathrm{F})\, \mathcal{K}(\Phi_{rad}) \left[\mathcal{O}(\mathbf{\Phi}_\mathrm{F},\Phi_{rad}) - \mathcal{O}(\mathbf{\Phi}_\mathrm{F})\right].\end{aligned} \tag{3.11}$$

Comparing equations (3.9) and (3.11), we see a similar structure: for moving from a PS approach at $\mathcal{O}(\alpha_s)$ to the correct NLO calculation, it is sufficient to perform the following replacements:

$$\mathrm{NLO}^{\mathrm{PS}} \leftrightarrow \mathrm{NLO}: \quad B(\mathbf{\Phi}_\mathrm{F}) \leftrightarrow \bar{B}(\mathbf{\Phi}_\mathrm{F}), \quad B(\mathbf{\Phi}_\mathrm{F})\, \mathcal{K}(\Phi_{rad}) \leftrightarrow R(\mathbf{\Phi}_\mathrm{FJ}). \tag{3.12}$$

We can thus construct the final POWHEG master formula as follows

$$\langle\mathcal{O}\rangle = \int \mathrm{d}\mathbf{\Phi}_\mathrm{F} \bar{B}(\mathbf{\Phi}_\mathrm{F}) \left\{ \Delta_\mathrm{pwg}(Q,Q_0)\, \mathcal{O}(\mathbf{\Phi}_\mathrm{F}) + \int_{Q_0}^{Q} \mathrm{d}\Phi_{rad} \frac{R(\mathbf{\Phi}_\mathrm{FJ})}{B(\mathbf{\Phi}_\mathrm{F})} \Delta_\mathrm{pwg}(Q,t)\, \mathcal{O}(\mathbf{\Phi}_\mathrm{F},\Phi_{rad}) \right\}, \tag{3.13}$$

where the POWHEG Sudakov form factor is defined as:

$$\Delta_\mathrm{pwg}(Q,Q_0) = \exp\left[ -\int_{Q_0}^{Q} \mathrm{d}\Phi_{rad} \frac{R(\mathbf{\Phi}_\mathrm{FJ})}{B(\mathbf{\Phi}_\mathrm{F})} \right] \equiv \Delta_\mathrm{pwg}(\mathbf{\Phi}_\mathrm{F},Q_0). \tag{3.14}$$

The last equivalence defines a new notation we will adopt in the next sections. To use the same notation as [5], we rename $Q_0$ as ${p_\mathrm{T}}^{\mathrm{min}}$. Correspondingly, the POWHEG differential cross section is obtained as

$$\mathrm{d}\sigma = \mathrm{d}\mathbf{\Phi}_\mathrm{F} \bar{B}(\mathbf{\Phi}_\mathrm{F}) \left\{ \Delta_\mathrm{pwg}(\mathbf{\Phi}_\mathrm{F}, {p_\mathrm{T}}^{\mathrm{min}}) + \int \mathrm{d}\Phi_{rad} \Delta_\mathrm{pwg}(\mathbf{\Phi}_\mathrm{F}, k_\mathrm{T}) \frac{R(\mathbf{\Phi}_\mathrm{FJ})}{B(\mathbf{\Phi}_\mathrm{F})} \right\}, \tag{3.15}$$

where ${p_\mathrm{T}}^{\mathrm{min}}$ is set to an IR cutoff (the default value is 0.89 GeV) and $k_\mathrm{T}$ is the transverse momentum of the radiation. The final cross section is thus the product of two terms: the first one is given by $\bar{B}$, which contains the fixed-order calculation at NLO accuracy; the second one is the term in brackets, which comes directly from the probabilistic description of the shower evolution, in which the spitting kernels are substituted with the correct matrix elements. The terms in brackets generate the first emission of the shower without using the soft and collinear approximation. If we generate events according to the differential cross section (3.15), we can apply the parton shower simply requiring that the transverse momentum of the shower emissions is smaller than the hardness of the radiation generate at LHE level.

We stress that:

| | **F** | **FJ** | **FJJ** | **F(≥3J)** |
|---|---|---|---|---|
| F @ POWHEG | NLO | LO | LL | LL |
| FJ @ POWHEG | - | NLO | LO | LL |
| FJ @ MiNLO′ | NLO | NLO | LO | LL |

Table 3.2.: Formal accuracy reached through the F, FJ and FJJ generators in the POWHEG framework according to the jet multiplicity. See text for more details.

- at large transverse momentum $k_{\mathrm{T}}$, $\sigma$ is the NLO cross section, up to NNLO corrections;
- we reproduce IR-safe observables at NLO, so that also the small $k_{\mathrm{T}}$ region is NLO accurate;
- the structure of the shower is untouched, as only the first emission (that is the hardest) is modified with the correct matrix element.

A more rigorous proof of the NLO accuracy of (3.15) is presented in [18].

## 3.2. The MiNLO′ method

In the previous section we discussed the main features of the POWHEG method, reviewing how to perform a matching between fixed-order calculations and parton showers at NLO accuracy in this framework. It is thus reasonable to ask ourselves whether it is possible to merge different POWHEG event generators with different jet multiplicities. More precisely, the POWHEG method can be used for studying the production of a colour singlet F at NLO+PS accuracy. In this F generator, the accuracy is NLO in distributions inclusive in F, LO in distributions with one jet and only LL (the accuracy of the shower) from the second jet onwards. It is also possible to construct a FJ generator, in which one jet J is already present at Born level. In this case, the NLO accuracy is retained in distributions inclusive in FJ, LO in FJJ and LL in FJJJ. This logic can be iterated to any number of jets at Born level. All these simulations present an overlap in the phase space and the accuracy they reach is complementary. It is thus natural to investigate the possibility of constructing a single framework in which the highest accuracy is retained in different jet bins. In this section we present the MiNLO′ method [40], which is an algorithm that merges the F and the FJ generators keeping the highest accuracy of both: the NLO accuracy is reached at the same time in distributions that are inclusive in F and FJ. See table 3.2 for a schematic representation.

The MiNLO′ method was proposed for merging different generators without the introduction of any unphysical scale. In fact, this merging problem was already addressed through different solutions [42, 43, 44, 45, 46, 47, 48, 49] in which typically an arbitrary scale is needed. This unphysical scale is used for partitioning the phase space according to the jet multiplicity: the result of each generator (F, FJ, FJJ, ...) contributes only to the jet bin for which it is NLO accurate. The different jet bins are then grouped together to obtain an inclusive result.

The presence of an unphysical parameter is clearly problematic: if this scale is too low, the obtained sample is dominated by high-multiplicity generators, while if it is too high, hard jets are described at low accuracy. For this reason, the dependence on this scale is deeply studied, as the final result cannot depend on unphysical parameters.

The MiNLO$'$ approach has been developed starting from the MiNLO method [39], which stands for *Multi-scale Improved NLO*. The MiNLO method was introduced as the NLO extension of the CKKW approach [50] in order to define an a priori criterion in the selection of central scales in NLO computations. We briefly summarize the main idea behind the method following the discussion presented in section 3 of [39].

We consider a generic production process at hadron colliders and we proceed as follows:

- We cluster the coloured partons in the event through the $k_\mathrm{T}$ algorithm, exactly like in the CKKW approach. The clustering should be consistent with the flavour of partons and, at each recombination, we assign the correct flavour to the newly defined particles. In this way, we can reconstruct the most likely branching history of the process. Once the skeleton of the event is known, we assign at each vertex $i$ $(i \in \{1, 2, ..., n\})$ a nodal scale $q_i$. These scales are set according to the CKKW procedure, which means that they are taken equal to the relative transverse momentum at which the clustering has been performed. When the clustering procedure is no longer possible, we remain with a set of particles that define the so-called *primary system*. We associate a scale $Q$ to the primary system equal to the invariant mass of the system itself. We recall that in the CKKW method a resolution scale $Q_0$ is introduced, defining the scale below which the cross section is inclusive. In the MiNLO approch, we set this scale to $Q_0 = q_1$.

- $n$ powers of the coupling $\alpha_s$ appearing in the Born, real and virtual corrections are set to $\mu_i = K_R\, q_i$, where $i$ runs over the $n$ vertices. The remaining $m$ powers of the strong coupling associated to the primary system are evaluated at $K_R\, Q$. $K_R$ is the usual scale variation factor.

- In the virtual correction there is the explicit appearance of the renormalization scale $\mu_R$ that is set as:

  $$\mu_R = (\mu_Q^m \times \prod_{i=1}^{n} \mu_i)^{1/(m+n)} \,. \tag{3.16}$$

  Moreover, the factorization scale $\mu_F$ is set to $\mu_F = K_F\, q_1$, where $K_F$ is the usual scale variation factor.

- We associate to each internal line connecting two vertices $i$ and $j$ (with $i$ closer to the primary vertex, meaning $q_i > q_j$) the following term

  $$\frac{\Delta(Q_0, q_i)}{\Delta(Q_0, q_j)}\,, \tag{3.17}$$

  where $\Delta$ is the Sudakov form factor:

  $$\Delta(Q_0, Q) = \exp\left\{ -\int_{Q_0}^{Q} \frac{\mathrm{d}q}{q} \left[ A(\alpha_s(q)) \log\frac{Q^2}{q^2} + B(\alpha_s(q)) \right\} \right. . \tag{3.18}$$

Since $Q_0 = q_1$, external lines directly connected to the first node $q_1$ are reweighted by $\Delta(Q_0, q_1) = 1$. The coefficients $A$ and $B$ admit an expansion in the strong coupling and can be calculated analytically through transverse momentum resummation [51, 52, 53]:

$$A(\alpha_s) = \left(\frac{\alpha_s}{2\pi}\right) A^{(1)} + \left(\frac{\alpha_s}{2\pi}\right)^2 A^{(2)} + \mathcal{O}\left(\frac{\alpha_s}{2\pi}\right)^3 , \tag{3.19}$$

$$B(\alpha_s) = \left(\frac{\alpha_s}{2\pi}\right) B^{(1)} + \mathcal{O}\left(\frac{\alpha_s}{2\pi}\right)^2 . \tag{3.20}$$

In the MiNLO approach, only $A^{(1)}$, $A^{(2)}$ and $B^{(1)}$ are encoded in the computation: this is a key point of the discussion, as it represents one of the main differences between the MiNLO and the MiNLO′ methods. The explicit expressions of the coefficients differ for quark and gluon lines. Note that we dropped a flavour index in (3.18) for keeping the notation easy. Defining $C_q = C_\mathrm{F} = 4/3$, $C_g = C_\mathrm{A} = 3$, the $A^{(i)}$ coefficients read

$$\begin{aligned} A^{(1)}_{q,g} &= 2\, C_{q,g} , \\ A^{(2)}_{q,g} &= 2\, C_{q,g} \left[ C_\mathrm{A} \left( \frac{67}{18} - \frac{\pi^2}{6} \right) - \frac{10}{9} T_\mathrm{R} n_f \right] , \end{aligned} \tag{3.21}$$

where $T_\mathrm{R} = 1/2$ and $n_f$ is the number of flavours. As for the $B^{(1)}$ coefficient, its expression is:

$$B^{(1)}_{q,g} = -2\gamma^{(1)}_{q,g} = \begin{cases} -3\, C_\mathrm{F} & \text{for } q\text{-lines} , \\ -4\,\pi\beta_0 & \text{for } g\text{-lines} . \end{cases} \tag{3.22}$$

- The inclusion of Sudakov form factors in the calculation introduces NLO corrections that spoil the accuracy of the computation. We thus need to subtract the first-order expansion $\mathcal{O}(\alpha_s)$ in the Born contribution. Schematically, we need to perform the following replacement

  $$B \to B \times \left(1 - \Delta^{(1)}\right) , \tag{3.23}$$

  where $B$ is the Born term and $\Delta^{(1)}$ is the first order expansion in $\alpha_s$ of the Sudakov form factor in (3.18).

- The last step consists in setting the remaining powers of the coupling in the real and virtual corrections to:

  $$\alpha_s = \frac{1}{n+m} \left( \sum_{i=1}^{n} \alpha_s(\mu_i) + m\, \alpha_s(\mu_Q) \right) . \tag{3.24}$$

  The idea behind this choice is the following: large QCD effects are associated with the nodal scales of the process and are thus related to the product of the coupling $\alpha_s$ at each scale and the Born term. The sum of these contributions thus corresponds to the sum of the values of the couplings at the scales involved in the process.

To summarize, the MiNLO method provides a prescription in the choice of the scales and the inclusion of Sudakov form factors for the NLO computation of processes involving jet production. A careful investigation of the accuracy of the method led the same authors to improve this methodology in the MiNLO′ formalism presented in [40]. Considering the production of a colour singlet F and defining an extra emission as J, the MiNLO′ method allows one to retain the NLO accuracy in both F and FJ at the same time. The modifications with respect to the MiNLO procedure are minimal but they allow us to merge generators with different jet multiplicities within the POWHEG framework keeping the highest accuracy of the two descriptions, as described at the beginning of this section.

We consider the production of a colour singlet F in association with one jet J, so that the usual POWHEG formula reads

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\mathbf{\Phi}_{\mathrm{FJ}}} = \bar{B}(\mathbf{\Phi}_{\mathrm{FJ}}) \left\{ \Delta_{\mathrm{pwg}}(\mathbf{\Phi}_{\mathrm{FJ}}, \Lambda_{\mathrm{pwg}}) + \int \mathrm{d}\Phi_{rad} \Delta_{\mathrm{pwg}}(\mathbf{\Phi}_{\mathrm{FJ}}, k_T) \frac{R(\mathbf{\Phi}_{\mathrm{FJJ}})}{B(\mathbf{\Phi}_{\mathrm{FJ}})} \right\}, \tag{3.25}$$

where the Born term $\bar{B}(\mathbf{\Phi}_{\mathrm{FJ}})$ is related to FJ production, while the terms in the curly brackets generate the extra radiation according to the POWHEG approach. $\Lambda_{\mathrm{pwg}}$ is the IR cutoff and $k_{\mathrm{T}}$ is the transverse momentum of the radiation. The MiNLO′ formalism modifies the $\bar{B}(\mathbf{\Phi}_{\mathrm{FJ}})$ term in this way:

- As done in the MiNLO method, we encode appropriate Sudakov form factors in the computation but the resummation coefficients that are now included are $A^{(1)}$, $A^{(2)}$, $B^{(1)}$ and $B^{(2)}$.
- The couplings entering the virtual and real contributions as well as in the first-order expansion of the Sudakov form factor are set to the transverse momentum of the colour singlet F. The same scale choice is performed for the factorization scale.

The POWHEG $\bar{B}(\mathbf{\Phi}_{\mathrm{FJ}})$ function thus reads

$$\bar{B}^{\mathrm{MiNLO}'}(\mathbf{\Phi}_{\mathrm{FJ}}) = e^{-\tilde{S}(p_{\mathrm{T}})} \left[ B(\mathbf{\Phi}_{\mathrm{FJ}}) \left( 1 + \frac{\alpha_s(p_{\mathrm{T}})}{2\pi} [\tilde{S}(p_{\mathrm{T}})]^{(1)} \right) + V(\mathbf{\Phi}_{\mathrm{FJ}}) + \int \mathrm{d}\Phi_{rad} R(\mathbf{\Phi}_{\mathrm{FJ}}, \Phi_{rad}) \right], \tag{3.26}$$

where we made contact with the MiNNLO$_{\mathrm{PS}}$ formalism introducing the following notation for the Sudakov form factor:

$$\tilde{S}(p_{\mathrm{T}}) = 2 \int_{p_{\mathrm{T}}}^{Q} \frac{\mathrm{d}q}{q} \left[ A\big(\alpha_s(q)\big) \ln \frac{Q^2}{q^2} + B\big(\alpha_s(q)\big) \right]. \tag{3.27}$$

The coefficients $A^{(1)}$, $A^{(2)}$ and $B^{(1)}$ are reported in equations (3.21) and (3.22), while the explicit expressions of the coefficients $B^{(2)}$ are [54, 53, 55]

$$\begin{aligned} B_q^{(2)} &= \left( -\frac{17}{12} - \frac{11\pi^2}{12} + 6\zeta_3 \right) C_{\mathrm{A}} C_{\mathrm{F}} + \left( -\frac{3}{4} + \pi^2 - 12\zeta_3 \right) C_{\mathrm{F}}^2 + \left( \frac{1}{6} + \frac{\pi^2}{6} \right) C_{\mathrm{F}} n_f \,, \\ B_g^{(2)} &= \left( \frac{11\zeta_2}{6} - 6\zeta_3 - \frac{16}{3} \right) C_{\mathrm{A}}^2 + \left( \frac{4}{3} - \frac{\zeta_2}{3} \right) C_{\mathrm{A}} n_f + n_f C_{\mathrm{F}} \,, \end{aligned} \tag{3.28}$$

where $\zeta_2 = 1.645...$ is the Riemann zeta function evaluated in $z = 2$. The term $[\tilde{S}(p_T)]^{(1)}$ in equation (3.26) is the first-order expansion of the Sudakov form factor defined in (3.27) and, as in the MiNLO procedure, it is needed not to spoil the NLO accuracy of the FJ computation. The inclusion of the $B^{(2)}$ coefficient together with setting the scales to the transverse momentum $p_T$ of F is crucial: it regularises the divergence in the Born level up to NLO accuracy in F because the singular part of the FJ cross section is encoded in the $B^{(2)}$ coefficients. A complete proof of the NLO accuracy of the method can be found in the original publication [40].

Using the MiNLO′ method it is possible to construct a NNLO+PS accurate event generator applying an a-posteriori reweighting [40]. More precisely, the weights of events generated using MiNLO′ can be rescaled using the following multi-differential reweighting factor:

$$w(\boldsymbol{\Phi}_F) = \frac{\left(\frac{d\sigma}{d\boldsymbol{\Phi}_F}\right)_{\text{NNLO}}}{\left(\frac{d\sigma}{d\boldsymbol{\Phi}_F}\right)_{\text{MiNLO}'}} . \tag{3.29}$$

By construction, distributions that are inclusive in F are NNLO accurate. Moreover, the NLO accuracy in FJ provided by the MiNLO′ method is not spoilt because we can expand the previous equation as

$$w(\boldsymbol{\Phi}_F) = \frac{c_0 + c_1\,\alpha_s + c_2\,\alpha_s^2}{c_0 + c_1\,\alpha_s + d_2\,\alpha_s^2} \approx 1 + \frac{c_2 - d_2}{c_0}\alpha_s^2 + \mathcal{O}(\alpha_s^3)\,, \tag{3.30}$$

which shows that the numerator and denominator start differing from order $\alpha_s^2$ so that the one-jet region is not spoilt by the reweighting procedure.

This method was applied to a few LHC processes (Higgs-boson production [56], Drell-Yan [57], $WH$ production [58], $ZH$ production [59], $WW$ production [60] and $H \to b\bar{b}$ decay [61]). However, a reweighting procedure becomes too challenging for complex processes: the more complicated the Born phase space, the more computationally demanding the procedure. For many processes of interest at the LHC this method is not feasible. In the next section we will present the MiNNLO$_{\text{PS}}$ method in which the same NNLO+PS accuracy is reached without requiring a CPU-exhausting reweighting.

## 3.3. The MiNNLO$_{\text{PS}}$ method

Reaching NNLO+PS accuracy is the current frontier for precise predictions at colliders and the MiNNLO$_{\text{PS}}$ framework [8] is a powerful tool for reaching this accuracy. Nowadays other methods are available, like the GENEVA framework [62, 47] and the UNNLOPS approach (applied only to simple processes like Drell-Yan [63] and Higgs production [64]), but in this context MiNNLO$_{\text{PS}}$ plays a special role thanks to its simplicity and flexibility. In the MiNNLO$_{\text{PS}}$ method, NNLO corrections are calculated on-the-fly during the event generation, which allows one to avoid a computationally intense reweighting. Moreover, no unphysical scales or boundaries are used for partitioning the phase space according to the jet multiplicity.

Last but not least, the matching with a $p_{\mathrm{T}}$-ordered shower is straightforward as MiNNLO$_{\mathrm{PS}}$ is embedded into the POWHEG method.

The MiNNLO$_{\mathrm{PS}}$ method was originally formulated for the production of a colour singlet in simple $2 \to 1$ processes [8], with an optimization of the method in [65]. It was later extended to a generic colour-singlet production [66] and finally to heavy-quark pair production [67]. Thanks to its simplicity, in the past few years the method was successfully applied to different processes: Higgs and $Z$ production [8], $Z\gamma$ production [66] (with the inclusion of anomalous couplings in [68]), $W^+W^-$ [69] and $ZZ$ [70], Higgsstrahlung with the decay of the Higgs boson into a pair of bottom quarks [2] (with an extension to SMEFT in [3]), $W^{\pm}Z$ including NLO EW effects [4], $t\bar{t}$ [67, 71] and $b\bar{b}$ [72] productions.

Schematically, we can summarize the MiNNLO$_{\mathrm{PS}}$ procedure in three main steps:

- Firstly, we generate a colour singlet F in association with a light parton J at NLO accuracy through the POWHEG method.
- The physical behaviour at low $p_{\mathrm{T}}$ is restored including an appropriate Sudakov form factor, so that the cross-section is finite even when the jet J becomes unresolved. Moreover, also higher-order corrections are included in order to obtain NNLO accuracy for distributions that are inclusive in F;
- Lastly, we generate exclusively the second radiated parton through the POWHEG method. Subsequent emissions are generated by the PS through the standard interface to POWHEG, meaning that these emissions are softer than the radiation generated at parton level in order to avoid double counting.

We now review in detail the construction of the MiNNLO$_{\mathrm{PS}}$ master formula following the discussion presented in the original publications [8, 65]. We consider the usual POWHEG formula for FJ production

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\mathbf{\Phi}_{\mathrm{FJ}}} = \bar{B}(\mathbf{\Phi}_{\mathrm{FJ}}) \left\{ \Delta_{\mathrm{pwg}}(\mathbf{\Phi}_{\mathrm{FJ}}, \Lambda_{\mathrm{pwg}}) + \int \mathrm{d}\Phi_{rad} \Delta_{\mathrm{pwg}}(\mathbf{\Phi}_{\mathrm{FJ}}, k_T) \frac{R(\mathbf{\Phi}_{\mathrm{FJJ}})}{B(\mathbf{\Phi}_{\mathrm{FJ}})} \right\}, \tag{3.31}$$

where $\bar{B}(\mathbf{\Phi}_{\mathrm{FJ}})$ already encodes the first emission J and the terms in the curly brackets generate the second radiation. Note that the Born contribution is divergent because of the presence of the light parton J, which gives rise to soft and collinear divergences that are not cancelled (no virtual contribution to F production is present).

The MiNNLO$_{\mathrm{PS}}$ method consists in modifying the term $\bar{B}(\mathbf{\Phi}_{\mathrm{FJ}})$ only, keeping the structure of the POWHEG calculation unaltered. Moreover, we need to encode in the computation NNLO accurate contributions. To do so, we consider the production of a colour singlet F with transverse momentum $p_{\mathrm{T}}$ and invariant mass $Q$. Using an all-order analytic formula [8], we can express the cross section differential in $\mathbf{\Phi}_{\mathrm{F}}$ and $p_{\mathrm{T}}$ as follows:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\mathrm{d}p_{\mathrm{T}}} = \frac{\mathrm{d}}{\mathrm{d}p_{\mathrm{T}}} \left\{ e^{-\tilde{S}(p_{\mathrm{T}})} \mathcal{L}(\mathbf{\Phi}_{\mathrm{F}}, p_{\mathrm{T}}) \right\} + R_f(\mathbf{\Phi}_{\mathrm{F}}, p_{\mathrm{T}}) \,. \tag{3.32}$$

This expression is derived from the analytic resummation of transverse momentum. The differential cross section is thus expressed as the sum of two contributions: the first one is the total derivative in $p_{\mathrm{T}}$ of an appropriate exponentiated Sudakov form factor $\tilde{S}(p_{\mathrm{T}})$ times a luminosity factor $\mathcal{L}(\mathbf{\Phi}_{\mathrm{F}}, p_{\mathrm{T}})$. This term is singular when $p_{\mathrm{T}} \to 0$. The second contribution $R_f(\mathbf{\Phi}_{\mathrm{F}}, p_{\mathrm{T}})$ is a finite term.

The Sudakov form factor is defined as

$$\tilde{S}(p_{\mathrm{T}}) = 2\int_{p_{\mathrm{T}}}^{Q} \frac{\mathrm{d}q}{q}\left[A\big(\alpha_s(q)\big)\ln\frac{Q^2}{q^2} + \tilde{B}\big(\alpha_s(q)\big)\right], \tag{3.33}$$

where the resummation coefficients $A(\alpha_s)$ and $B(\alpha_s)$ are considered up to the following perturbative order:

$$A(\alpha_s) = \left(\frac{\alpha_s}{2\pi}\right)A^{(1)} + \left(\frac{\alpha_s}{2\pi}\right)^2 A^{(2)} + \left(\frac{\alpha_s}{2\pi}\right)^3 A^{(3)}, \tag{3.34}$$

$$\tilde{B}(\alpha_s) = \left(\frac{\alpha_s}{2\pi}\right)B^{(1)} + \left(\frac{\alpha_s}{2\pi}\right)^2 \tilde{B}^{(2)}. \tag{3.35}$$

The explicit expressions of the $A^{(i)}$ and $B^{(i)}$ coefficients can be found in [51, 52, 53]. More precisely, $A^{(1)}$ and $A^{(2)}$ are defined in equation (3.21) while $A^{(3)}$ is

$$\begin{aligned} A^{(3)}_{q,g} = 2\,C_{q,g}\Bigg\{ C_{\mathrm{A}}^2\left[\frac{245}{24} - \frac{67}{9}\frac{\pi^2}{6} + \frac{11}{6}\zeta_3 + \frac{11}{5}\left(\frac{\pi^2}{6}\right)^2\right] + C_{\mathrm{F}}n_f\left[-\frac{55}{24} + 2\zeta_3\right] \\ + C_{\mathrm{A}}n_f\left[-\frac{209}{108} + \frac{10}{9}\frac{\pi^2}{6} - \frac{7}{3}\zeta_3\right] + n_f^2\left[-\frac{1}{27}\right]\Bigg\}, \end{aligned} \tag{3.36}$$

where $\zeta_3 = 1.202...$ is the Riemann zeta function $\zeta(z)$ evaluated in $z = 3$. As for the $B^{(i)}$ coefficients, $B^{(1)}$ is defined in (3.22) while $\tilde{B}^{(2)}$ is defined as follows:

$$\tilde{B}^{(2)}_{q,g} = B^{(2)}_{q,g} + 2\zeta_3(A^{(1)})^2 + 2\pi\beta_0 H^{(1)}. \tag{3.37}$$

The term $H^{(1)}$ represents the one-loop hard function and $B^{(2)}$ is defined in equation (3.28). Notice that $\tilde{B}^{(2)}_{q,g}$ depends explicitly on the specific hard process through the hard virtual correction $H^{(1)}$.

The luminosity factor $\mathcal{L}(\mathbf{\Phi}_{\mathrm{F}}, p_{\mathrm{T}})$ is defined as

$$\begin{aligned} \mathcal{L}(\mathbf{\Phi}_{\mathrm{F}}, p_{\mathrm{T}}) = \sum_{c,c'} \frac{\mathrm{d}B_{c,c'}(\mathbf{\Phi}_{\mathrm{F}})}{\mathrm{d}\mathbf{\Phi}_{\mathrm{F}}} \sum_{i,j}\Big\{ (\tilde{C}^{[a]}_{ci} \otimes f^{[a]}_i)\,\tilde{H}(p_{\mathrm{T}})\,(\tilde{C}^{[b]}_{c'j} \otimes f^{[b]}_j) \\ + (G^{[a]}_{ci} \otimes f^{[a]}_i)\,\tilde{H}(p_{\mathrm{T}})\,(G^{[b]}_{c'j} \otimes f^{[b]}_j)\Big\}, \end{aligned} \tag{3.38}$$

where $[a]$ and $[b]$ represent the initial legs and $c, c', i, j$ are flavour indices. $B$ is the Born contribution for F production, $\tilde{C}$ and $G$ are collinear coefficients up to $\mathcal{O}(\alpha_s^2)$ (see section 4 in [8] for their complete expressions), and $\tilde{H}$ encodes hard virtual corrections up to two loops as follows:

$$\begin{aligned} \tilde{H} &= 1 + \frac{\alpha_s}{2\pi}H^{(1)} + \left(\frac{\alpha_s}{2\pi}\right)^2 \tilde{H}^{(2)} \\ \tilde{H}^{(2)} &= H^{(2)} + 2\zeta_3 A^{(1)} B^{(1)}. \end{aligned} \tag{3.39}$$

In order to make contact with the MiNLO′ formalism, we can rewrite expression (3.32) as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\mathrm{d}p_{\mathrm{T}}} = e^{-\tilde{S}(p_{\mathrm{T}})}\left\{D(p_{\mathrm{T}}) + \frac{R_f(p_{\mathrm{T}})}{e^{-\tilde{S}(p_{\mathrm{T}})}}\right\}, \tag{3.40}$$

where we have introduced the key ingredient of the MiNNLO$_{\mathrm{PS}}$ procedure:

$$D(p_{\mathrm{T}}) \equiv -\frac{\mathrm{d}\tilde{S}(p_{\mathrm{T}})}{\mathrm{d}p_{\mathrm{T}}}\mathcal{L}(p_{\mathrm{T}}) + \frac{\mathrm{d}\mathcal{L}(p_{\mathrm{T}})}{\mathrm{d}p_{\mathrm{T}}}. \tag{3.41}$$

The finite contribution $R_f(p_{\mathrm{T}})$ can be rewritten as the difference of the full NLO calculation and the singular contribution expandend up to the second order in the coupling

$$R_f(p_{\mathrm{T}}) = \frac{\mathrm{d}\sigma_{\mathrm{FJ}}^{(\mathrm{NLO})}}{\mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\mathrm{d}p_{\mathrm{T}}} - \frac{\alpha_s(p_{\mathrm{T}})}{2\pi}\left[\frac{\mathrm{d}\sigma^{\mathrm{sing}}}{\mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\mathrm{d}p_{\mathrm{T}}}\right]^{(1)} - \left(\frac{\alpha_s(p_{\mathrm{T}})}{2\pi}\right)^2\left[\frac{\mathrm{d}\sigma^{\mathrm{sing}}}{\mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\mathrm{d}p_{\mathrm{T}}}\right]^{(2)}, \tag{3.42}$$

where the notation $[X]^{(i)}$ stands for the $i$-th term in the perturbative expansion of $X$. The NLO cross section for FJ production is:

$$\frac{\mathrm{d}\sigma_{\mathrm{FJ}}^{(\mathrm{NLO})}}{\mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\mathrm{d}p_{\mathrm{T}}} = \frac{\alpha_s(p_{\mathrm{T}})}{2\pi}\left[\frac{\mathrm{d}\sigma_{\mathrm{FJ}}}{\mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\mathrm{d}p_{\mathrm{T}}}\right]^{(1)} - \left(\frac{\alpha_s(p_{\mathrm{T}})}{2\pi}\right)^2\left[\frac{\mathrm{d}\sigma_{\mathrm{FJ}}}{\mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\mathrm{d}p_{\mathrm{T}}}\right]^{(2)}. \tag{3.43}$$

In order to preserve the perturbative accuracy of equation (3.40), we expand the curly brackets in powers of the coupling using the following power counting formula

$$\int_{\Lambda}^{Q} \mathrm{d}p_{\mathrm{T}}\frac{1}{p_{\mathrm{T}}}\alpha_s^m(p_{\mathrm{T}})\ln\frac{Q}{p_{\mathrm{T}}}\,e^{-\tilde{S}(p_{\mathrm{T}})} \approx \mathcal{O}\left(\alpha_s^{m-\frac{n+1}{2}}(Q)\right), \tag{3.44}$$

which means that we retain NNLO accuracy if we keep contributions up to $\mathcal{O}(\alpha_s^3)$.

Combining equations (3.40), (3.42), (3.43) and using the power counting formula (3.44), we obtain the final NNLO differential cross section:

$$\begin{aligned}\frac{\mathrm{d}\sigma}{\mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\mathrm{d}p_{\mathrm{T}}} = e^{-\tilde{S}(p_{\mathrm{T}})}\Bigg\{&\frac{\alpha_s(p_{\mathrm{T}})}{2\pi}\left[\frac{\mathrm{d}\sigma_{\mathrm{FJ}}}{\mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\mathrm{d}p_{\mathrm{T}}}\right]^{(1)}\left(1+\frac{\alpha_s(p_{\mathrm{T}})}{2\pi}[\tilde{S}(p_{\mathrm{T}})]^{(1)}\right) + \left(\frac{\alpha_s(p_{\mathrm{T}})}{2\pi}\right)^2\left[\frac{\mathrm{d}\sigma_{\mathrm{FJ}}}{\mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\mathrm{d}p_{\mathrm{T}}}\right]^{(2)} \\ &+ \left(\frac{\alpha_s(p_T)}{2\pi}\right)^3\left(D(p_T) - \frac{\alpha_s(p_T)}{2\pi}[D(p_T)]^{(1)} - \left(\frac{\alpha_s(p_T)}{2\pi}\right)^2[D(p_T)]^{(2)}\right)F^{corr}(\mathbf{\Phi}_{\mathrm{FJ}}) + \mathrm{regular}\Bigg\}.\end{aligned} \tag{3.45}$$

We thus recover exactly the MiNLO′ formula (3.26), with the inclusion of extra terms through which NNLO accuracy is reached. We can thus redefine the POWHEG $\bar{B}$ function encoding all the ingredients obtained in the previous equation, thus obtaining a NNLO+PS accurate event generator.

The regular contributions can be neglected as they are formally N$^3$LO accurate: they arise from the $\mathcal{O}(\alpha_s^3)$ expansion of the finite term $R_f$ in equation (3.32), which is free from $1/p_{\mathrm{T}}$

singularities and vanishes when $p_{\mathrm{T}}$ goes to zero. Moreover, note that it is possible to use the following truncation

$$D(p_{\mathrm{T}}) - \frac{\alpha_s(p_{\mathrm{T}})}{2\pi}[D(p_{\mathrm{T}})]^{(1)} - \left(\frac{\alpha_s(p_{\mathrm{T}})}{2\pi}\right)^2 [D(p_{\mathrm{T}})]^{(2)} = \left(\frac{\alpha_s(p_{\mathrm{T}})}{2\pi}\right)^3 [D(p_{\mathrm{T}})]^{(3)} + \mathcal{O}(\alpha_s^4(p_{\mathrm{T}}))\,, \tag{3.46}$$

as it was done in the first formulation of the MINNLO$_{\mathrm{PS}}$ method [8]. In [65] it was shown that the impact of this truncation might be numerically not negligible in specific processes and in kinematic configurations with small $p_{\mathrm{T}}$. In fact, the truncation presented in (3.46) introduces a different treatment of subleading terms when compared to fixed-order NNLO calculations. For this reason, in the default formulation of the method the truncation is not performed anymore and all the terms of $\mathcal{O}(\alpha_s^4)$ and higher are kept. In this way, the total derivative in (3.32) can be exactly reconstructed.

The term $F^{corr}$ that multiplies the $D(p_{\mathrm{T}})$ terms in equation (3.45) is needed for spreading the NNLO corrections in the FJ phase space. More precisely, all the MINLO$'$ terms are directly related to the phase space of F with one jet ($\mathbf{\Phi}_{\mathrm{FJ}}$) or with two jets ($\mathbf{\Phi}_{\mathrm{FJJ}}$). By contrast, the NNLO $D(p_{\mathrm{T}})$ contributions are obtained from a resummed calculation in the $p_{\mathrm{T}} \to 0$ limit and they thus depend on the phase space of the colour singlet only ($\mathbf{\Phi}_{\mathrm{F}}$) in which the information of the extra radiation has been integrated out. The $D(p_{\mathrm{T}})$ terms explicitly depend on the $p_{\mathrm{T}}$ of the colour singlet but this dependence is not associated with a well-defined phase space point of the full kinematics (either $\mathbf{\Phi}_{\mathrm{FJ}}$ or $\mathbf{\Phi}_{\mathrm{FJJ}}$). In fact, the presence of a non-zero $p_{\mathrm{T}}$ requires a recoil of the colour singlet against at least one parton, but we do not have any information on such a parton.

Concretely, solving this issue means defining a method for associating each value of $D(p_{\mathrm{T}})$ with a specific phase space point of $\mathbf{\Phi}_{\mathrm{FJ}}$. We thus introduce a mapping that projects $\mathbf{\Phi}_{\mathrm{FJ}}$ to $\mathbf{\Phi}_{\mathrm{F}}$ smoothly in the limit $p_{\mathrm{T}} \to 0$. This procedure goes under the name of *spreading* because the $D(p_{\mathrm{T}})$ contributions are spread over the $\mathbf{\Phi}_{\mathrm{FJ}}$ phase space in a way that allows one to reproduce (3.45) after integration.

There are different possibilities for defining the spreading, either uniformly or through a suitable distribution. We consider the following spreading factor

$$F_\ell^{corr}(\mathbf{\Phi}_{\mathrm{FJ}}) = \frac{\mathcal{I}_\ell(\mathbf{\Phi}_{\mathrm{FJ}})}{\sum_{\ell'} \int \mathrm{d}\mathbf{\Phi}'_{\mathrm{FJ}}\, \mathcal{I}_{\ell'}\, \delta(p_{\mathrm{T}} - p_{\mathrm{T}}{}')\delta(\mathbf{\Phi}_{\mathrm{F}} - \mathbf{\Phi}'_{\mathrm{F}})}\,, \tag{3.47}$$

where $\mathbf{\Phi}_{\mathrm{F}}$ ($\mathbf{\Phi}'_{\mathrm{F}}$) is obtained projecting $\mathbf{\Phi}_{\mathrm{FJ}}$ ($\mathbf{\Phi}'_{\mathrm{FJ}}$) into the F phase space (e.g. according to the FKS mapping [21, 22] that preserves the rapidity of the colour singlet F. Explicit expressions for the $\mathbf{\Phi}_{\mathrm{FJ}} \to \mathbf{\Phi}_{\mathrm{F}}$ mapping can be found in appendix A of [8].). $\mathcal{I}_\ell$ is an arbitrary function of $\mathbf{\Phi}_{\mathrm{FJ}}$ and $\ell$ defines the flavour structure of the FJ process. With $p_{\mathrm{T}}$ we indicate the transverse momentum of the colour singlet in the $\mathbf{\Phi}_{\mathrm{FJ}}$ phase space. Taken an arbitrary function $G(\mathbf{\Phi}_{\mathrm{F}}, p_{\mathrm{T}})$,

$F_\ell^{corr}$ should satisfy:

$$\sum_\ell \int \mathrm{d}\mathbf{\Phi}'_{\mathrm{FJ}}\, G(\mathbf{\Phi}'_{\mathrm{F}}, p_{\mathrm{T}}')\, F_\ell^{corr}(\mathbf{\Phi}'_{\mathrm{FJ}}) = \int \mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\, \mathrm{d}p_{\mathrm{T}}\, G(\mathbf{\Phi}_{\mathrm{F}}, p_{\mathrm{T}})$$

$$\times \sum_\ell \int \mathrm{d}\mathbf{\Phi}'_{\mathrm{FJ}}\, \delta(\mathbf{\Phi}_{\mathrm{F}} - \mathbf{\Phi}'_{\mathrm{F}})\, \delta(p_{\mathrm{T}} - p_{\mathrm{T}}')\, F_\ell^{corr}(\mathbf{\Phi}'_{\mathrm{FJ}}) = \int \mathrm{d}\mathbf{\Phi}_{\mathrm{F}}\, \mathrm{d}p_{\mathrm{T}}\, G(\mathbf{\Phi}_{\mathrm{F}}, p_{\mathrm{T}})\,. \quad (3.48)$$

The function $\mathcal{I}_\ell$ can be taken equal to one, thus spreading the NNLO corrections in a flat way over the phase space:

$$\mathcal{I}_\ell(\mathbf{\Phi}_{\mathrm{FJ}}) = 1\,. \quad (3.49)$$

For this trivial choice, the integral in equation (3.47) can be analytically solved. However, this spreading gives rise to spurious effects when the jet is produced at large rapidities. One can thus adopt a more physical distribution, in which the spreading depends on the physical matrix element

$$\mathcal{I}_\ell(\mathbf{\Phi}_{\mathrm{FJ}}) = |M_\ell^{\mathrm{FJ}}(\mathbf{\Phi}_{\mathrm{FJ}})|^2\, (f^{[a]} f^{[b]})_\ell\,, \quad (3.50)$$

where $|M_\ell^{\mathrm{FJ}}(\mathbf{\Phi}_{\mathrm{FJ}})|^2$ is the squared Born matrix element for FJ production and $(f^{[a]} f^{[b]})_\ell$ is the product of the parton densities for a specific flavour configuration $\ell$. In this way, the $D(p_{\mathrm{T}})$ terms are spread according to the rapidity distribution of the radiation. Even though this choice has a correct physical interpretation, the calculation can become computationally intense for complex processes. An approximated version of equation (3.50) is thus preferable: we can consider the squared Born matrix element in the collinear limit

$$|M_\ell^{\mathrm{FJ}}(\mathbf{\Phi}_{\mathrm{FJ}})|^2 \simeq |M_\ell^{\mathrm{F}}(\mathbf{\Phi}_{\mathrm{F}})|^2 P_\ell(\Phi_{rad})\,, \quad (3.51)$$

where $P_\ell(\Phi_{rad})$ is the collinear splitting function. Noting that the squared matrix element $|M_\ell^{\mathrm{F}}(\mathbf{\Phi}_{\mathrm{F}})|^2$ for F production cancels exactly in equation (3.47), the spreading function $\mathcal{I}_\ell(\mathbf{\Phi}_{\mathrm{FJ}})$ is simply set to:

$$\mathcal{I}_\ell(\mathbf{\Phi}_{\mathrm{FJ}}) = P_\ell(\Phi_{rad})(f^{[a]} f^{[b]})_\ell\,. \quad (3.52)$$

The computation of (3.47) is thus much faster and complex processes can be handled as well.

We now analyse the procedure used for switching off the Sudakov form factor and the $D(p_T)$ terms at high $p_T$, which is needed for not introducing spurious contributions when $p_{\mathrm{T}}$ becomes large. In fact, in this region the MiNNLO$_{\mathrm{PS}}$ formula differs from the NLO FJ cross section for terms beyond accuracy only (starting at $\mathcal{O}(\alpha_s^3)$). The choice of the exact procedure has some arbitrariness, provided the logarithmic structure of the computation at small $p_{\mathrm{T}}$ is preserved. We can introduce modified logarithms of the form

$$\ln \frac{Q}{p_{\mathrm{T}}} \to L \equiv \frac{1}{p} \ln\left(1 + \left(\frac{Q}{p_{\mathrm{T}}}\right)^p\right), \quad (3.53)$$

where $p$ is a free positive parameter that regulates how fast the logarithms go to zero. The larger $p$, the faster the logarithms tend to 0 at large $p_{\mathrm{T}}$ values. Note that the formal accuracy of the MiNNLO$_{\mathrm{PS}}$ method is not spoilt because in the limit $p_{\mathrm{T}} \to 0$ the calculation remains unaffected. This modification of the logarithms introduces terms beyond accuracy in equation (3.45), and it must be performed at the level of (3.32) in order to retain the total cross section. To do so, we need the following adjustments:

- The renormalization and factorization scales in $D(p_T)$ are set to
$$\mu_R = K_R\, Q\, e^{-L}, \quad \mu_F = K_R\, Q\, e^{-L}\,, \tag{3.54}$$
where $K_R$ and $K_F$ are scale variation factors varied between 1/2 and 2.
- The lower bound of the integral in the Sudakov form factor (3.33) is set to:
$$p_\mathrm{T} \to Q\, e^{-L}\,. \tag{3.55}$$
- The $D(p_\mathrm{T})$ terms are supplemented with an appropriate Jacobian factor that directly depends on the modified logarithms:
$$D(p_\mathrm{T}) \to \mathcal{J}_Q\, D(p_\mathrm{T}), \quad \mathcal{J}_Q \equiv \frac{(Q/p_\mathrm{T})^p}{1+(Q/p_\mathrm{T})^p}\,. \tag{3.56}$$

We note that equation (3.45) reproduces the correct NNLO result only if it can be evaluated down to small values of $p_\mathrm{T}$. More precisely, the integration in $p_\mathrm{T}$ in (3.32) should be carried out using a low integration bound, where the integrand tends to zero because of the suppression of the Sudakov form factor. Approaching small values of $p_\mathrm{T}$ can be problematic because of the internal cutoff $\Lambda_\mathrm{PDF}$ of PDFs: each set of PDFs has an intrinsic IR cutoff $\Lambda_\mathrm{PDF}$ related to the low scale for which a fit on data is no longer possible (because data are not available in the very low energy region). Below this IR cutoff, the PDF is set to zero. In order to approach small $p_\mathrm{T}$ values, this truncation of the PDFs at $\Lambda_\mathrm{PDF}$ should be avoided and a consistent DGLAP evolution for lower scales should be applied.

The strategy adopted in the MiNNLO$_\mathrm{PS}$ method is described in [8] and it proceeds as follows: PDF grids are read from the LHAPDF [73] package and then used for constructing corresponding HOPPET [74] grids. The advantage of the HOPPET package is the fast evaluation of convolutions with coefficient functions. At scales $\mu_F > \Lambda_\mathrm{PDF}$ HOPPET grids are equal to the LHAPDF ones, while for values $\mu_F < \Lambda_\mathrm{PDF}$ we perform a DGLAP evolution freezing the number of active flavours to the one provided by LHAPDF at $\mu_F = \Lambda_\mathrm{PDF}$. Using this prescription one can perform a running of the PDFs to small $p_\mathrm{T} \sim \mu_F$ values.

The non-perturbative regime at low $p_\mathrm{T}$ can be smoothly approached with the introduction of a damping factor $g(p_\mathrm{T})$
$$\mu_R = K_R\left(Q\, e^{-L} + Q_0\, g(p_\mathrm{T})\right), \quad \mu_F = K_F\left(Q\, e^{-L} + Q_0\, g(p_\mathrm{T})\right), \tag{3.57}$$
where $Q_0$ is a non-perturbative parameter that regularises the Landau singularity. The choice of $g(p_\mathrm{T})$ is to a large extent arbitrary: once the analytic form of the damping factor is chosen, the same scale settings are adopted in the Sudakov form factor in equation (3.33). Moreover, an appropriate Jacobian factor $\mathcal{J}_{Q_0}$ that depends on $g(p_\mathrm{T})$ should multiply the $D(p_\mathrm{T})$ terms.

We remind the reader that there is some freedom in the choice of the scales in the NLO FJ cross section: while scale settings are constrained by $p_\mathrm{T}$ resummation in the $p_\mathrm{T} \to 0$ limit, for finite values of the transverse momentum the MiNNLO$_\mathrm{PS}$ method allows for some arbitrariness. The default scale choice is
$$\mu_R = K_R\, p_\mathrm{T}\,, \quad \mu_F = K_F\, p_\mathrm{T}\,, \tag{3.58}$$

or, when the damping is introduced:

$$\mu_R = K_R\,(p_\mathrm{T} + g(p_\mathrm{T})\,Q_0)\,, \quad \mu_F = K_F(p_\mathrm{T} + g(p_\mathrm{T})\,Q_0)\,. \tag{3.59}$$

In this way, the choice of the scales is consistent with the scale used in the $D(p_\mathrm{T})$ terms and in the Sudakov form factor when $p_\mathrm{T}$ becomes small, reproducing a correct matching in this limit. However, the scales of the NLO computation can also be set as in equations (3.54) and (3.57), such that when $p_\mathrm{T}$ becomes large, scales are of the order of the invariant mass $Q$ of the colour singlet.

To summarize: the MiNNLO$_\mathrm{PS}$ method provides a powerful framework for matching NNLO calculations and parton showers. All the terms needed to reach this accuracy are already present at the level of the event generation, thus providing an efficient tool that does not require any a-posteriori reweighting. The matching with the parton shower is straightforward because it follows the POWHEG approach. In Part II of this thesis, we will present some cutting-edge phenomenological results obtained using this method. We first analyse Higgsstrahlung with subsequent decay of the Higgs boson to a pair of bottom quarks in chapter 4. We then present an extension of $ZH$ production with $H \to b\bar{b}$ decay in the SM Effective Field Theory in chapter 5. We conclude in chapter 6 with a detailed analysis of the combination of QCD and EW corrections matched to QCD and/or QED parton showers for $W^{\pm}Z$ production.

# Part II.
# PRECISION LHC PHENOMENOLOGY

# 4. Higgsstrahlung with $H \to b\bar{b}$ decay at NNLO+PS accuracy in the SM

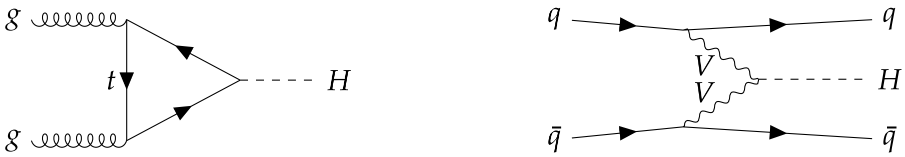
## 4.1. Motivation

The Higgs boson is a unique particle in the Standard Model because it is related to fundamental properties of nature and how the world appears to us. In fact, the Higgs field has a vacuum expectation value different from zero that is directly linked to the masses of the elementary particles of the SM. However, the SM mechanism for the generation of masses is not the only one possible: improving our description of key processes that involve the Higgs boson can be the path towards the discovery of new physics.

In the last decade, extraordinary theoretical and experimental achievements were reached in the investigation of the Higgs sector, but this exploration is still at an early stage. Given that a better understanding of the Higgs sector can possibly lead to an extension of the SM, the Higgs boson is and, even more so, will be at the core of the rich LHC physics programme. In this context, precise theoretical calculations are crucial in order to detect a small deviation in the data/theory comparison.

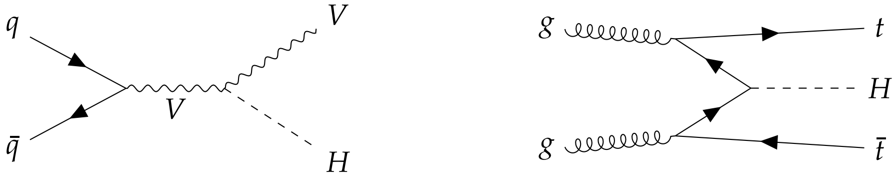
A detailed review of the Higgs boson properties and state-of-the-art measurements and predictions can be found in the latest PDG [75]. At the LHC, the Higgs boson is produced through the four main channels represented in figure 4.1.

The dominant production mode is gluon-gluon fusion (panel (a)), which is characterised by a heavy-quark loop. Since the Yukawa coupling of a particle is proportional to its mass, the leading contribution is given by a loop with top quarks. Even if this process is loop-induced and thus suppressed by powers of the strong coupling, this channel is strongly enhanced by the large gluon PDF and the large top Yukawa coupling. Indeed, gluon-gluon fusion constitutes 90% of the total inclusive cross section for Higgs production at the LHC. The drawback of this channel is the large QCD background which renders experimental measurements extremely challenging. For this reason, other Higgs production channels which allow for a cleaner experimental signature are indispensable for precise tests of the Higgs sector. The second-largest production mode is vector-boson fusion (panel (b)), in which the Higgs boson couples to two vector bosons emitted by the incoming partons. The typical feature of this process is the presence of two back-to-back hard jets given by the scattered partons in the forward and backward regions. The inclusive cross section for this channel is around 7% of the total production rate of the Higgs boson. The next most relevant production channel is the associated production of the Higgs boson with a vector boson $V$, also referred to as *Higgsstrahlung* (panel (c)). This production mode has a small cross section (∼2% of the total inclusive cross section for Higgs production) but it provides a clean experimental



(a) Gluon-gluon fusion

(b) Vector-boson fusion



(c) Higgsstrahlung

(d) Associated $t\bar{t}$ production

Figure 4.1.: Main production channels of the Higgs boson at the LHC. Panel (a) is gluon-gluon fusion; panel (b) is vector-boson fusion; panel (c) is Higgsstrahlung; panel (d) is associated production with a pair of top quarks.

signature. In order to increase the production rate, it is possible to consider the decay of the Higgs boson into a pair of bottom quarks, which constitutes the largest branching fraction (roughly 60%). Despite the large $H \to b\bar{b}$ decay rate, it had been thought for a long time that this channel could not be observed at the LHC because of the large background until it was proposed to consider a boosted regime in which the Higgs boson has a large transverse momentum and to apply substructure techniques for the reconstruction of the bottom quarks [76]. The last production mode for the Higgs boson is the associated production with a couple of top quarks (panel (d)). The production rate associated to this mode is low ($\sim$1% of the total Higgs production rate) and this channel is hard to detect because of the large background.

This chapter is dedicated to precise predictions for the production of a Higgs boson through Higgstrahlung, with subsequent decay of the Higgs boson into a pair of bottom quarks. Sample Feynman diagrams can be found in figure 4.2. This process is of fundamental importance for precision physics in the Higgs sector as it provides direct access to the bottom Yukawa coupling and it actually yields the highest sensitivity to its determination.

Both Higgsstrahlung and $H \to b\bar{b}$ decay were extensively studied in the past. As far as $VH$ production is concerned, a first matching of NLO QCD calculations with parton showers was presented in [77] within the POWHEG framework. In this work, the MiNLO$'$ method [40] was used. A combination at NLO accuracy of both QCD and EW computations matched to parton showers is presented in [78]. The loop induced $gg \to HZ$ process has been matched to PS at LO in [79, 80]. At NNLO+PS accuracy, the GENEVA method [62] was applied to $VH$ production in [81] while the MiNLO$'$+ reweighting procedure described in section 3.2 was applied in [59, 58]. As for the $H \to b\bar{b}$ decay, NNLO QCD results in the massless approximation are available in [82, 83, 84] and a matching with parton showers at NNLO+PS accuracy is available in [61, 85, 86].

In this chapter, we present results for $W^{\pm}H$ and $ZH$ production processes followed by

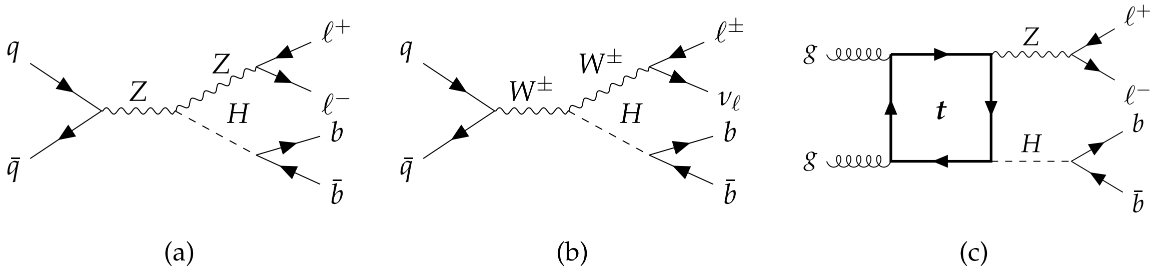


Figure 4.2.: Feynman diagrams for $VH$ production with subsequent $H \to b\bar{b}$ decay. Panels (a) and (b) are $q\bar{q}$-induced (Drell-Yan like) channels at LO, while panel (c) is a $gg$-induced channel that enters the calculation at NNLO.

$H \to b\bar{b}$ decay at NNLO+PS accuracy in both production and decay. These results were originally presented by the author and collaborators in [2]. The chapter is organised as follows: we first describe our calculation in section 4.2 and we validate our implementation in section 4.3. Our phenomenological results are presented in section 4.4, in which we show predictions for $W^+H$ and $ZH$. The corresponding results for $W^-H$ are presented in Appendix A. We also present a comparison with data in the same section. We conclude with a detailed analysis of the impact of the jet-clustering algorithm used for the reconstruction of $b$-jets in section 4.5.

## 4.2. Outline of the calculation

We consider the Higgsstrahlung process for both neutral and charged vector bosons

$$pp \to ZH \to \ell^+\ell^-H, \quad pp \to ZH \to \nu_\ell\bar{\nu}_\ell H \quad \text{and} \quad pp \to W^\pm H \to \ell^\pm\nu_\ell H\,, \tag{4.1}$$

with the subsequent decay of the Higgs boson $H$ into a pair of bottom quarks:

$$H \to b\bar{b}\,. \tag{4.2}$$

With $\ell$ we indicate a massless lepton ($\ell \in \{e, \mu, \tau\}$), while $\nu_\ell$ ($\bar{\nu}_\ell$) is the corresponding (anti-) neutrino.

In figures 4.2 (a) and 4.2 (b) we show sample Feynman diagrams of the process at hand in the $q\bar{q}$-induced channel at LO. Figure 4.2 (c) represents the $gg$-induced channel, which is mediated by a heavy-quark loop. This channel enters the calculation at NNLO accuracy and, due to charge conservation, it exists only for $ZH$ production. This contribution has been neglected in our phenomenological results in section 4.4.3, but it is included when presenting total cross sections in section 4.4.2 and in the comparison with data in section 4.4.4.

Our predictions are NNLO+PS accurate in both the production and the decay of the Higgs boson. In order to reach this accuracy, we adopt the following strategy: we first generate separately events associated to the production and the decay stages of the Higgs boson at

NNLO accuracy. For the production stage $pp \to VH$, we employ the MiNNLO$_{\text{PS}}$ method, while for the decay $H \to b\bar{b}$ we use the MiNLO$'$+ reweighting approach, which will be referred to as NNLOPS method from now on. We stress that the MiNNLO$_{\text{PS}}$ method could have been used also for the generation of decay events, but the application of the NNLOPS approach is straightforward for the $H \to b\bar{b}$ decay (more details are given later in section 4.2.2). The production and decay events are then combined using a narrow-width approximation for the Higgs boson. Using this procedure, production and decay stages are considered completely factorized and we thus neglect interference terms between final-state bottom quarks and initial-state partons. This approximation is motivated by the fact that interference terms are exactly zero at $\mathcal{O}(\alpha_s)$, as the Higgs boson is a colour singlet, and they are strongly suppressed by the small width of the Higgs ($\Gamma_H \simeq 4.1\,\text{MeV}$) at $\mathcal{O}(\alpha_s^2)$. Once the events for the complete process are obtained, we consistently shower them using PYTHIA8 [87].

### 4.2.1. Generation of production events

Our generators for the production stage $pp \to VH$ of the Higgs boson are implemented in both the POWHEG-BOX-V2 [41] and the POWHEG-BOX-RES [88] frameworks. The vector boson $V = W^{\pm}, Z$ can decay in any lepton channel. Our POWHEG-BOX-V2 implementation relies on the $VH$+jet generators developed in [77], while our code in POWHEG-BOX-RES is based on the $VH$+jet generators presented in [78]. The two implementations are fully compatible within numerical uncertainties.

We remind the reader that we do not include in the computation $gg$-induced processes (panel (c) in figure 4.2) except in section 4.4.2 for the calculation of total cross sections and in section 4.4.4 for the comparison with data. In these cases, we generate LO+PS accurate results for $gg$-induced processes in the POWHEG-BOX-V2 framework and we combine them to the $q\bar{q}$-initiated channels in an additive way at the level of differential distributions. Moreover, there are other corrections to both $WH$ and $ZH$ productions mediated by a heavy-quark loop in which the Higgs boson is radiated by the heavy quark. A detailed review of these contributions can be found in [89], where they were calculated for the first time. Using the notation of [89], in our implementation we neglect virtual contributions $V_{\text{I}}$ and $V_{\text{II}}$ and include real contributions $R_{\text{I}}$ and $R_{\text{II}}$ (see the introduction section of [89]). Note that $R_{\text{II}}$ contributes to $ZH$ production only. Both $V_{\text{I,II}}$ and $R_{\text{I,II}}$ contribute to the total cross section to about 0.6%.

In our implementation, the matrix elements are taken from different providers: the Born (i.e. $VH$+jet) and real (i.e. $VH$+2 jets) matrix elements are obtained from the interface [90] between POWHEG and MadGraph v4 [91]. The virtual amplitude for $VH$+jet is calculated analytically [92]. The one-loop and two-loop amplitudes for $VH$ production are obtained by encoding the correct rescaling in the quark vertex function. For the calculation of the $D$ terms in the MiNNLO$_{\text{PS}}$ formula (3.45), we use HOPPET [74] for a fast evaluation of the convolutions with PDFs. Furthermore, we compute polylogarithms in the collinear coefficient functions using hplog [93].

To avoid spurious higher-order logarithmic terms in the MiNNLO$_{\text{PS}}$ formula (3.45), we adopt a modified logarithm (3.53) that smoothly turns off NNLO corrections in the region $p_{\text{T}} > m_{VH}/2$, where $p_{\text{T}}$ and $m_{VH}$ are the transverse momentum and the invariant mass of the $VH$ colour singlet. The hard scale in the logarithms is consistently changed to $m_{VH}/2$ [71]. At high $p_{\mathrm{T}}$, we set the scales of the calculation through the `largeptscales 1` option so that equation (3.58) is used. At small $p_{\mathrm{T}}$, the default MiNNLO$_{\mathrm{PS}}$ prescription in the choice of the scales is employed. Furthermore, we set $Q_0 = 0$ GeV in (3.57) and we freeze the strong coupling and the PDFs at 0.8 GeV. We also switch on the POWHEG option doublefsr 1 [94], so that $q \to qg$ and $g \to q\bar{q}$ splittings are treated symmetrically for the definition of the starting scale of the shower. We use PYTHIA8 [87] as parton shower and we adopt default settings, in particular for the recoil scheme.

### 4.2.2. Generation of decay events

As for the generation of decay events, we follow the procedure described in section 2.2 of [61] and we use the same implementation within the POWHEG-BOX-V2 framework. NNLO+PS accuracy is reached through the NNLOPS method, which requires an a-posteriori reweighting. Even though the MiNNLO$_{\mathrm{PS}}$ method could have been applied, the reweighting procedure is straightforward in this case as the reweighting factor is given by a number only, which is the NNLO decay width of the Higgs boson decaying into $b$-quarks:

$$\mathcal{W}(\mathbf{\Phi}_{b\bar{b}}) = \frac{\Gamma^{\mathrm{NNLO}}_{H\to b\bar{b}}}{\Gamma^{\mathrm{MiNLO}'}_{H\to b\bar{b}}}\,. \tag{4.3}$$

We smoothly switch off the reweighting factor in hard events that are already NLO accurate using a modified version of (4.3), as explained in [61]. We consider the three-jet resolution parameter $y_3$ in the Cambridge algorithm [95, 96], which defines the separation between two-jet and three-jet events, and we construct the following reweighting factor:

$$\mathcal{W}(\mathbf{\Phi}_{b\bar{b}}) = h(y_3)\,\frac{\Gamma^{\mathrm{NNLO}}_{H\to b\bar{b}} - \Gamma^{\mathrm{MiNLO}',B}_{H\to b\bar{b}}}{\Gamma^{\mathrm{MiNLO}',A}_{H\to b\bar{b}}} + (1 - h(y_3))\,, \tag{4.4}$$

where

$$\begin{aligned}
\Gamma^{\mathrm{MiNLO}',A}_{H\to b\bar{b}} &= \int \mathrm{d}\mathbf{\Phi}_{b\bar{b}g}\left(\frac{\Gamma^{\mathrm{MiNLO}'}_{H\to b\bar{b}}}{\mathrm{d}\mathbf{\Phi}_{b\bar{b}g}}\,h(y_3)\right),\\
\Gamma^{\mathrm{MiNLO}',B}_{H\to b\bar{b}} &= \int \mathrm{d}\mathbf{\Phi}_{b\bar{b}g}\left(\frac{\Gamma^{\mathrm{MiNLO}'}_{H\to b\bar{b}}}{\mathrm{d}\mathbf{\Phi}_{b\bar{b}g}}\,(1 - h(y_3))\right).
\end{aligned} \tag{4.5}$$

The $h(y_3)$ function is defined as

$$h(y_3) = \frac{1}{1 + \left(\frac{y_3}{y_{3,ref}}\right)^{y_{3,pow}}}\,, \tag{4.6}$$

where the suggested values for the two parameters are $y_{3,ref} = e^{-4}$ and $y_{3,pow} = 2$, as explained in [61]. By construction, $h(y_3)$ tends to one when soft radiation is produced, while

$h(y_3) \ll 1$ when a hard jet is produced. Using this procedure, the $\mathcal{W}(\mathbf{\Phi}_{b\bar{b}})$ factor is applied only to processes without hard radiation.

All the amplitudes needed for the computation of the decay events have been taken from [84].

### 4.2.3. Combination of production and decay events

Once both production and decay events are produced, we combine them using the procedure described in [61]. The first step consists in substituting the Higgs boson generated in a production event with its decay products taken from a decay event. This requires a reshuffling of the momenta of the decay products, to restore energy-momentum conservation. Small off-shell effects suppressed by the Higgs width $\Gamma_H$ are neglected (these terms are of order $\mathcal{O}(\Gamma_H/m_H)$, where $m_H$ is the Higgs boson mass). Furthermore, the decay events are generated in the Higgs rest frame while the production events are in the laboratory frame, thus an appropriate boost is needed.

We then modify the weights of the combined events $w_{\text{full}}$ using this narrow-width approximation formula

$$w_{\text{full}} = \frac{w_{\text{prod}} \cdot w_{\text{dec}}}{\Gamma_H} = w_{\text{prod}} \cdot \text{Br}(H \to b\bar{b}) \cdot \frac{w_{\text{dec}}}{\Gamma_{H\to b\bar{b}}}, \tag{4.7}$$

where the weights for the production and the decay events are indicated as $w_{\text{prod}}$ and $w_{\text{dec}}$, respectively. $\text{Br}(H \to b\bar{b})$ represents the branching fraction of the Higgs boson to bottom quarks (in our code, it is taken as an input), while $\Gamma_{H\to b\bar{b}}$ is the $H \to b\bar{b}$ decay width directly computed from the decay events. By construction, if we integrate inclusively the last factor in equation (4.7), we obtain one: the total cross section for the full event is the product of the cross section for the production and the branching ratio of the Higgs boson decaying into a pair of bottom quarks, regardless of the accuracy of the last factor.

The specific value of the bottom Yukawa coupling used in the calculation is irrelevant, as it cancels exactly in the ratio in the last factor in equation (4.7).

The matching of the full event with the parton shower is non-trivial as the starting scales for the production and the decay stages of the process are different. We follow the procedure described in [61] also in this case. In the combined event file, we store the starting scale of the production stage (`scalup`), while the corresponding value for the decay (`scalup`$_{\texttt{dec}}$) is recomputed on-the-fly starting from the decay kinematics. We adopt the usual definition of hardness for final-state radiation used in POWHEG [97]. `scalup` is read by the shower from the Les Houches Event file and it is automatically used as the starting scale for all the splittings coming from the production stage. As for the decay, we let the shower radiate in the entire kinematically allowed phase space. We then apply an a-posteriori veto to check that the hardness of the splittings generated by the shower are consistent with the emission produced by POWHEG from the Higgs decay at Les Houches Event level. The hardness of a splitting is defined as

$$t = 2\,(p_{\text{rad}} \cdot p_{\text{em}})\,\frac{E_{\text{rad}}}{E_{\text{em}}}, \tag{4.8}$$

where $E_{\text{rad}}$ ($E_{\text{em}}$) and $p_{\text{rad}}$ ($p_{\text{em}}$) are the energy and 4-momentum of the radiated (emitting)

| | $\sigma$ [fb] |
|---|---|
| MiNNLO$_{\text{PS}}$ | $96.72(4)^{+1.9\%}_{-0.6\%}$ |
| NNLOPS | $96.69(3)^{+1.3\%}_{-1.3\%}$ |

Table 4.1.: Inclusive cross sections for $pp \to e^+\nu_e H$ production using the MiNNLO$_{\text{PS}}$ and NNLOPS methods.

particle. An event is accepted only if the hardness of all the splittings generated by the shower in the Higgs decay is smaller than `scalup`$_{\texttt{dec}}$. If this condition is not met, we reject the event and we try to shower it again till the requirement is fulfilled. After 1000 failed attempts, we reject the event.

## 4.3. Validation of the calculation

In this section we validate our MiNNLO$_{\text{PS}}$ implementation against NNLOPS results. NNLOPS results for $W^\pm H$ production are taken from [58], while $ZH$ production with subsequent decay of the Higgs boson into bottom quarks is presented in [61]. In the former case, the Higgs boson is kept stable, so that the full calculation for $W^\pm H$ production with $H \to b\bar{b}$ decay at NNLO+PS accuracy is obtained with the MiNNLO$_{\text{PS}}$ method in [2] for the first time.

### 4.3.1. $W^\pm H$ production

We consider $W^\pm H$ production with a stable Higgs boson, and we compare our MiNNLO$_{\text{PS}}$ predictions [2] with NNLOPS results [58]. Since only $W^+H$ is presented in [58], we present a comparison for the positively charged vector boson only.

We study proton-proton collisions at 13 TeV centre-of-mass energy for the production of $e^+\nu_e H$. In order to present a reliable comparison, the input parameters are taken from [58]: $m_W = 80.399\,\text{GeV}$, $\Gamma_W = 2.085\,\text{GeV}$, $m_H = 125.0\,\text{GeV}$. The electroweak parameters read: $\alpha = 1/132.3489$ and $\sin^2\theta_W = 0.2226$. The Higgs boson is considered on-shell and contributions where the Higgs boson is radiated from a heavy-quark loop ($V_{\text{I,II}}$ and $R_{\text{I,II}}$) are neglected, as done in [58]. The PDF set is MMHT2014nnlo68cl, for which the corresponding running coupling is $\alpha_s(m_Z^2) = 0.118$.

The renormalization and factorization scales are set according to the MiNNLO$_{\text{PS}}$ procedure and error bars are obtained through 7-point scale variation, where $\mu_R$ and $\mu_F$ are varied by a factor of two above and below the central scale with the constraint $1/2 \leq \mu_R/\mu_F \leq 2$. We present results for an inclusive setup, without any fiducial cut.

Table 4.1 shows the inclusive cross sections obtained with the MiNNLO$_{\text{PS}}$ and the NNLOPS methods. Results are in perfect agreement at the sub-permille level, within numerical errors. The MiNNLO$_{\text{PS}}$ results have slightly asymmetric error bars when compared to the NNLOPS ones. This is not unexpected because of the different scale settings and treatment of higher-order terms in the two calculations.

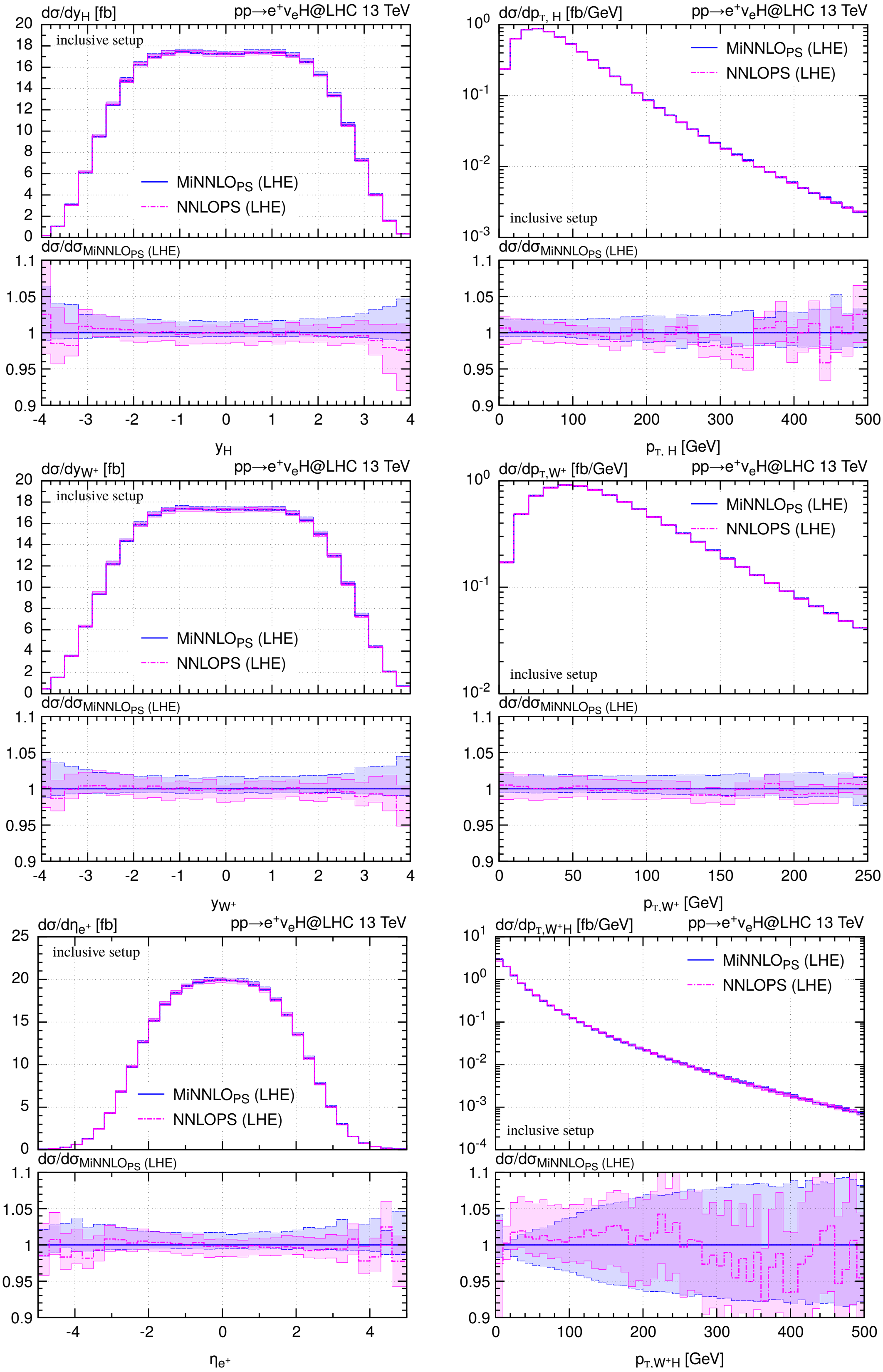


Figure 4.3.: Differential distributions for $e^+\nu_e H$ production obtained with the MiNNLO$_{\text{PS}}$ and the NNLOPS methods at LHE level.

| | $\sigma$ [fb] |
|---|---|
| POWHEG-BOX-RES | 26.51(1) |
| POWHEG-BOX-V2 | 26.63(9) |
| MCFM | 26.567(1) |

Table 4.2.: Inclusive cross sections for $pp \to e^+e^-H$ production obtained through two independent MiNNLO$_{\text{PS}}$ implementations within POWHEG-BOX-RES and POWHEG-BOX-V2, and MCFM. Digits in brackets refer to statistical error only.

Figure 4.3 shows a comparison of the MiNNLO$_{\text{PS}}$ and the NNLOPS predictions for different differential distributions at LHE level. More precisely, we show the rapidity ($y_H$) and transverse momentum ($p_{\text{T},H}$) of the Higgs boson, the rapidity ($y_{W^+}$) and the transverse momentum ($p_{\text{T},W^+}$) of the $W^+$ boson, the rapidity of the positron ($\eta_{e^+}$) and the transverse momentum of the $W^+H$ system ($p_{\text{T},W^+H}$).

All rapidity distributions, which are NNLO accurate observables, are in excellent agreement. Discrepancies between the two computations are at $\sim 1\%$ level, with small statical fluctuations in the tails. The asymmetry in the error bars of the MiNNLO$_{\text{PS}}$ results compared to the NNLOPS ones is visible. A similar conclusion holds also for the transverse momenta of the $H$ and $W^+$ bosons, where the agreement in the central value is within $1-2\%$. As for the $p_T$ of the $W^+H$ system, the difference between MiNNLO$_{\text{PS}}$ and NNLOPS is slightly larger, up to $3-4\%$: this distribution is effectively only NLO accurate at $p_{\text{T},W^+H} \gg 0$ and there are matching ambiguities in the intermediate region, which is reflected in the enlarged error bars in this region. Apart from this small difference, we see a good agreement as the two calculations are fully compatible within uncertainties.

We now comment on the size of the error bars of the two calculations in the $p_{\text{T},W^+H}$ distribution, as we observe a difference in the range $30\,\text{GeV} \lesssim p_{\text{T},W^+H} \lesssim 90\,\text{GeV}$. This discrepancy is related to the different way in which NNLO corrections are spread in the entire spectrum. In fact, in the MiNNLO$_{\text{PS}}$ approach we use modified logarithms while the NNLOPS calculation requires a $p_{\text{T}}$-dependent multi-differential reweighting. We stress that the shape of the MiNNLO$_{\text{PS}}$ error bars is closer to the one of a resummed calculation matched to a fixed-order computation, as pointed out in [8].

The comparison of MiNNLO$_{\text{PS}}$ and NNLOPS results for $e^+\nu_eH$ production shows a very good agreement between the two calculations. This validates our MiNNLO$_{\text{PS}}$ implementation.

### 4.3.2. $ZH$ production with $H \to b\bar{b}$ decay

The validation of our calculation for $ZH$ production with subsequent decay of the Higgs boson into a couple of $b$-quarks is carried out in two steps. Firstly, we validate the production stage only, considering $pp \to ZH$ with a stable Higgs boson.

We consider our default MiNNLO$_{\text{PS}}$ implementation in the POWHEG-BOX-RES framework, a completely independent MiNNLO$_{\text{PS}}$ implementation in POWHEG-BOX-V2 and a fixed-order result obtained with MCFM [98]. The comparison is performed in an inclusive setup at LHE level. Our integrated cross sections are presented in table 4.2, where we see a remarkable agreement among the three independent calculations.

We then include the decay of the Higgs boson into our computation and we compare the full $pp \to ZH \to e^+e^-b\bar{b}$ MiNNLO$_{\text{PS}}$ result to [61]. Also in this case, the validation is performed at LHE level. All the input parameters and cuts are taken from [61]. The PDF set is PDF4LHC15_nnlo_mc, which corresponds to $\alpha_s(m_Z^2)$ =0.118. Our SM parameters are:

$$\begin{aligned} G_F &= 1.166387 \times 10^{-5}\,\text{GeV}^{-2}\,, & \Gamma_Z &= 2.4952\,\text{GeV}\,,\\ m_W &= 80.398\,\text{GeV}\,, & m_H &= 125.0\,\text{GeV}\,,\\ \Gamma_W &= 2.141\,\text{GeV}\,, & \Gamma_H &= 4.14\,\text{MeV}\,.\\ m_Z &= 91.1876\,\text{GeV}\,, \end{aligned}$$

The EW coupling and the weak mixing angle are defined as:

$$\alpha = \sqrt{2} G_F m_W^2 (1 - m_W^2/m_Z^2)/\pi\,, \qquad \cos^2\theta_W = m_W^2/m_Z^2\,. \tag{4.9}$$

We include contributions $R_{\text{I,II}}$ in which the Higgs boson is radiated from a heavy-quark loop (see section 4.2.1). We consider both top and bottom loops and the corresponding pole masses are $m_b$ = 4.92 GeV and $m_t$ = 173.2 GeV. As far as the $H \to b\bar{b}$ implementation is concerned, we evaluate the bottom Yukawa coupling in the $\overline{\text{MS}}$ scheme at a scale equal to the Higgs mass $m_H$, corresponding to $y_b(m_H)$ = 1.280 $\times 10^{-2}$. We recall that the specific value of $y_b$ does not affect the calculation because it cancels exactly in the combination of production and decay events as shown in equation (4.7). The branching ratio for $H \to b\bar{b}$ decay is $\text{Br}(H \to b\bar{b})$ = 0.5824. In the production stage, the choice of the renormalization and factorization scales follows the MiNNLO$_{\text{PS}}$ prescription. In the decay, the renormalization scale is set to the Higgs boson mass $m_H$. Scale uncertainties are obtained through 7-point scale variation with the constraint $1/2 \le \mu_R/\mu_F \le 2$. We correlate the scale variation factors in production and decay.

| fiducial cuts for $pp \to ZH \to e^+e^-b\bar{b}$ |
|---|
| $p_{\text{T},e} > 7\,\text{GeV}$, $p_{\text{T},e_1} > 27\,\text{GeV}$, $\lvert\eta_e\rvert < 2.5$ |
| $81\,\text{GeV} < m_{e^+e^-} < 101\,\text{GeV}$ |
| $\geq$2 $b$-jets (flavour-$k_T$ [99], R=0.4) |
| $p_{\text{T},b\text{-jet}} > 27\,\text{GeV}$, $\lvert\eta_{b\text{-jet}}\rvert < 2.5$ |

Table 4.3.: Fiducial phase space for $e^+e^-b\bar{b}$ production as defined in [61].

In this comparison, we apply the fiducial cuts presented in table 4.3. In the jet-clustering procedure, if the clustering algorithm identifies more than two $b$-jets, we consider the pair whose invariant mass is closer to the Higgs-boson mass.

Differential cross sections are shown in table 4.4. The agreement between MiNNLO$_{\text{PS}}$ and NNLOPS is good (with a 1.4% difference) and fully covered by scale uncertainties. The

| | $\sigma$ [fb] |
|---|---|
| MiNNLO$_{\text{PS}}$ | $6.261(7)^{+0.9\%}_{-1.8\%}$ |
| NNLOPS | $6.348(6)^{+1.2\%}_{-1.4\%}$ |

Table 4.4.: Fiducial cross sections for $pp \to ZH \to e^+e^-b\bar{b}$ production using the MiNNLO$_{\text{PS}}$ and NNLOPS methods.

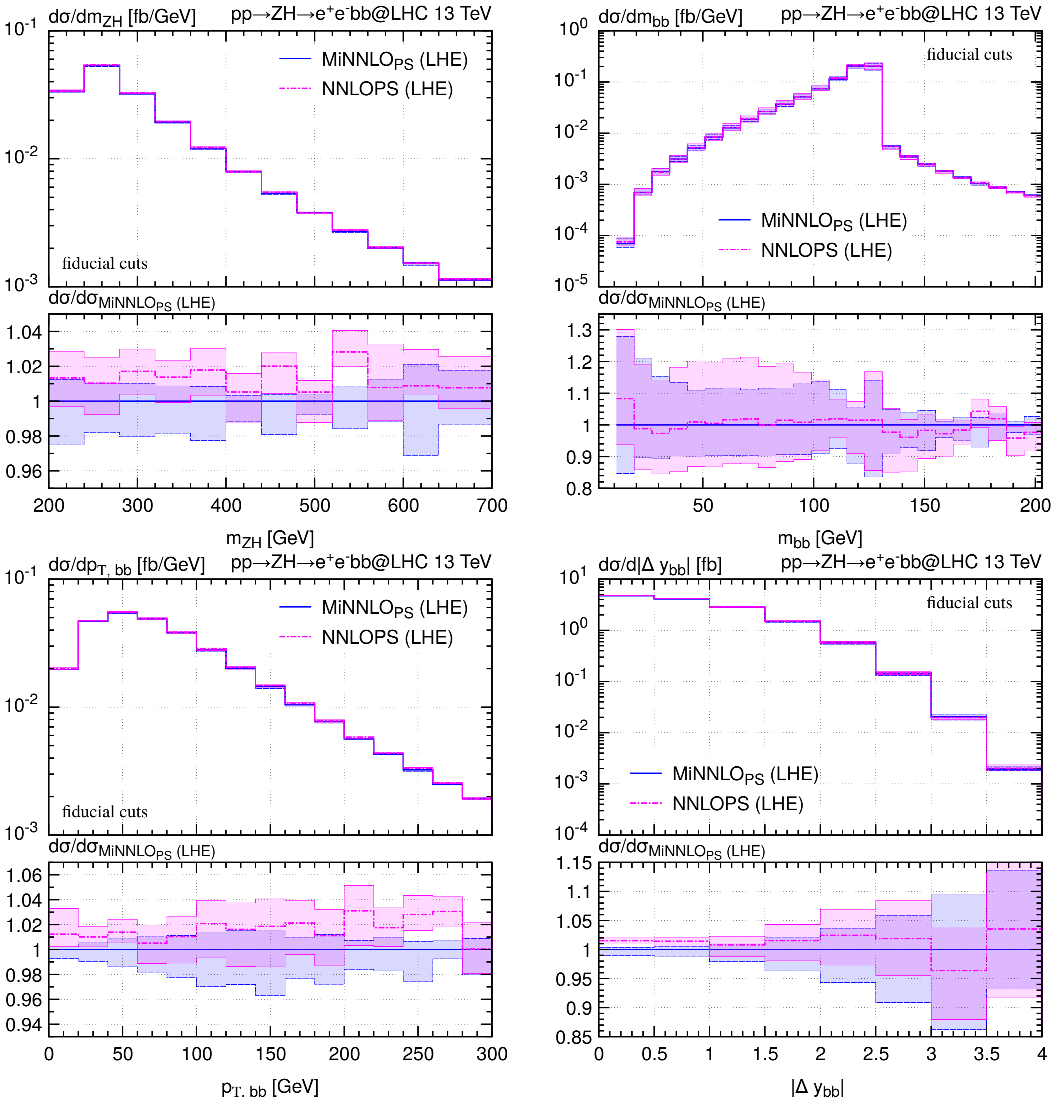


Figure 4.4.: Differential ditributions for $e^+e^-b\bar{b}$ production obtained with the MiNNLO$_{\text{PS}}$ and the NNLOPS methods at LHE level in the fiducial setup reported in table 4.3.

discrepancy between the two results is slightly larger than when comparing inclusive cross sections, because of the more exclusive setup.

Differential observables are presented in figures 4.4. In particular, we show the invariant mass of the $ZH$ system ($m_{ZH}$), the invariant mass ($m_{bb}$) and transverse momentum ($p_{\mathrm{T},bb}$) of the two reconstructed $b$-jets and the difference in rapidity between the two $b$-jets ($|\Delta y_{b,b}|$). We observe a good agreement between the two calculations, with differences that are smaller than 1-2% overall, with slightly larger statistical fluctuations in the tails. We have studied many other differential observables and for all of them the level of agreement between the two calculations is good. This comparison thus validates our MiNNLO$_{\mathrm{PS}}$ calculation.

## 4.4. Phenomenological results

In this section we present our differential predictions for $pp \to e^\pm \nu_e b\bar{b}$ and $pp \to e^+e^- b\bar{b}$ showing both MiNLO$'$ and MiNNLO$_{\mathrm{PS}}$ results. We present cross sections in both inclusive and fiducial setups in 4.4.2 and we discuss differential distributions in 4.4.3. Our MiNNLO$_{\mathrm{PS}}$ prediction is compared to ATLAS data [100] in 4.4.4. We also analyse the impact of different jet-clustering algorithms in section 4.5.2.

### 4.4.1. Input and parameters

Our results are obtained using the NNPDF31_nnlo_as_0118 [101] PDF set, which corresponds to $\alpha_s(m_Z^2)$=0.118. Our SM parameters are taken from the PDG [9], namely:

$$\begin{aligned} G_F &= 1.166379 \times 10^{-5}\,\mathrm{GeV}^{-2}\,, & \Gamma_Z &= 2.4952\,\mathrm{GeV}\,,\\ m_W &= 80.379\,\mathrm{GeV}\,, & m_H &= 125.09\,\mathrm{GeV}\,,\\ \Gamma_W &= 2.085\,\mathrm{GeV}\,, & \Gamma_H &= 4.1\,\mathrm{MeV}\,.\\ m_Z &= 91.1876\,\mathrm{GeV}\,, \end{aligned}$$

The EW coupling and the weak mixing angle are defined as:

$$\alpha = \sqrt{2} G_F m_W^2 (1 - m_W^2/m_Z^2)/\pi\,, \qquad \cos^2\theta_W = m_W^2/m_Z^2\,. \tag{4.10}$$

The $H \to b\bar{b}$ branching ratio is $\mathrm{Br}(H \to b\bar{b}) = 0.5824$. In $pp \to ZH \to e^+e^- b\bar{b}$ we include the $R_{\mathrm{I,II}}$ contributions, in which the Higgs boson is radiated from a heavy quark loop: the pole mass of the bottom quark is set to $m_b = 4.78\,\mathrm{GeV}$ and the pole mass of the top quark to $m_t = 172.5\,\mathrm{GeV}$. The bottom Yukawa coupling is $y_b(m_H)$ = 1.280 $\times 10^{-2}$: as already pointed out, our final result does not depend on this value because it cancels out exactly when interfacing production and decay events (see equation 4.7).

The parton shower is generated by PYTHIA8 [87] with the Monash tune [102]. In our interface to PYTHIA8, all the settings are kept to the standard ones, in particular for the recoil scheme. Effects from hadronization, underlying event and QED shower are not taken into account. We include hadronization effects in section 4.4.4 only, so that the comparison with data is more reliable.

| `fiducial-YR` [103] | |
|---|---|
| $pp \to W^\pm H \to e^\pm \nu_e b\bar{b}$ | $pp \to ZH \to e^+e^- b\bar{b}$ |
| $p_{\mathrm{T},e} > 15\,\mathrm{GeV}$, $p_{\mathrm{T,miss}} > 15\,\mathrm{GeV}$, $\lvert\eta_e\rvert < 2.5$ | $p_{\mathrm{T},e} > 15\,\mathrm{GeV}$, $\lvert\eta_e\rvert < 2.5$ |
| | $75\,GeV < m_{e^+e^-} < 105\,\mathrm{GeV}$ |
| $\geq$2 $b$-jets (anti-$k_T$ [104], R=0.4) | $\geq$2 $b$-jets (anti-$k_T$ [104], R=0.4) |
| $p_{\mathrm{T},b\text{-jet}} > 25\,\mathrm{GeV}$, $\lvert\eta_{b\text{-jet}}\rvert < 2.5$ | $p_{\mathrm{T},b\text{-jet}} > 25\,\mathrm{GeV}$, $\lvert\eta_{b\text{-jet}}\rvert < 2.5$ |
| `fiducial-ATLAS` [100] | |
| $pp \to W^\pm H \to \ell^\pm \nu_\ell b\bar{b}$ | $pp \to ZH \to \ell^+\ell^- b\bar{b}/\nu_\ell\bar{\nu}_\ell b\bar{b}$ |
| $\lvert y_H\rvert < 2.5$ | $\lvert y_H\rvert < 2.5$ |
| categories: | categories: |
| $p_{\mathrm{T},W} \in [250{,}400]\,\mathrm{GeV}$, | $p_{\mathrm{T},Z} \in [250{,}400]\,\mathrm{GeV}$, |
| $p_{\mathrm{T},W} \in [400, \infty]\,\mathrm{GeV}$ | $p_{\mathrm{T},Z} \in [400, \infty]\,\mathrm{GeV}$ |

Table 4.5.: Fiducial phase space in the `fiducial-YR` and the `fiducial-ATLAS` setups.

In the next sections, we present our results in three different setups: an inclusive setup (`inclusive`), a fiducial setup inspired by the CERN Yellow Report [103] (`fiducial-YR`) and a fiducial setup taken from [100] (`fiducial-ATLAS`). These setups are summarized in table 4.5.

### 4.4.2. Inclusive and fiducial cross sections

In this section we present integrated cross sections in both `inclusive` and `fiducial-YR` setups. Table 4.6 shows our MiNLO′ and MiNNLO$_{\mathrm{PS}}$ predictions for both $WH$ and $ZH$ productions with $H \to b\bar{b}$ decay. In the $ZH$ case, we report separately the results with and without $gg$-induced contributions.

The MiNNLO$_{\mathrm{PS}}$ corrections are positive and roughly 5-6% with respect to MiNLO′ calculation in both setups, if we do not consider the loop-induced $gg$-contribution for $ZH$ production. As expected, we observe a significant reduction of the error bars when moving from the MiNLO′ to the MiNNLO$_{\mathrm{PS}}$ calculation (errors are reduced by more than a factor of two). Considering only the inclusive setup, MiNLO′ and MiNNLO$_{\mathrm{PS}}$ predictions are compatible within scale variations. By contrast, in the `fiducial-YR` setup, the MiNNLO$_{\mathrm{PS}}$ corrections are generally not covered by scale variations: we verified that the MiNNLO$_{\mathrm{PS}}$ results is covered by the MiNLO′ error bars if we uncorrelate the scales in the combination of production and decay events.

Considering the $gg \to ZH$ contribution, its effect is quite large because it represents about 8% in the inclusive and 11% in the fiducial phase space of the MiNNLO$_{\mathrm{PS}}$ cross sections. As it is only LO accurate, the error bars coming from loop-induced $gg$-diagrams are the dominant uncertainties.

$$pp \to W^+H \to e^+\nu_e b\bar{b}$$

| $\sigma$ [fb] | `inclusive` | `fiducial-YR` |
|---|---|---|
| MiNLO$'$ | $54.04^{+6.6\%}_{-3.6\%}$ | $20.13^{+2.3\%}_{-3.1\%}$ |
| MiNNLO$_{\text{PS}}$ | $57.44^{+1.7\%}_{-0.8\%}$ | $21.27^{+1.3\%}_{-1.3\%}$ |

$$pp \to W^-H \to e^-\bar{\nu}_e b\bar{b}$$

| $\sigma$ [fb] | `inclusive` | `fiducial-YR` |
|---|---|---|
| MiNLO$'$ | $33.82^{+6.6\%}_{-3.6\%}$ | $13.07^{+2.4\%}_{-3.3\%}$ |
| MiNNLO$_{\text{PS}}$ | $35.87^{+1.5\%}_{-0.7\%}$ | $13.77^{+1.5\%}_{-1.6\%}$ |

$$pp \to ZH \to e^+e^- b\bar{b}$$

| $\sigma$ [fb] | `inclusive` | `fiducial-YR` |
|---|---|---|
| MiNLO$'$ | $14.88^{+6.7\%}_{-3.7\%}$ | $5.21^{+2.2\%}_{-3.0\%}$ |
| MiNNLO$_{\text{PS}}$ (no $gg \to ZH$) | $15.79^{+1.8\%}_{-0.9\%}$ | $5.48^{+1.2\%}_{-1.2\%}$ |
| MiNNLO$_{\text{PS}}$ (with $gg \to ZH$) | $16.99^{+3.6\%}_{-2.3\%}$ | $6.07^{+3.4\%}_{-2.9\%}$ |

Table 4.6.: Integrated cross sections for $W^{\pm}H$ and $ZH$ production with $H \to b\bar{b}$ decay in the `inclusive` and the `fiducial-YR` setups.

### 4.4.3. Differential distributions

In this section we show our phenomenological results considering differential distributions in the `fiducial-YR` phase space defined in table 4.5. We present our MiNNLO$_{\text{PS}}$ results for different observables for $W^+H$ and $ZH$ production[1]. In particular, we show: the invariant mass ($m_{bb}$) and transverse momentum ($p_{\mathrm{T},bb}$) of the bottom-quark pair, the absolute difference in rapidity between the two bottom quarks ($|\Delta y_{b,b}|$), the transverse momentum of the hardest bottom quark ($p_{\mathrm{T},b_{\mathrm{hard}}}$) and of the positron/electron ($p_{\mathrm{T},e^+}/p_{\mathrm{T},e^-}$), and the invariant mass of the colour-singlet system ($m_{WH}/m_{ZH}$). In this context, with *bottom-quark pair* we denote the two jets with at least one bottom quark whose invariant mass is closest to the Higgs-boson mass.

In our plots, the MiNNLO$_{\text{PS}}$ result is the blue and solid curve while the MiNLO$'$ result is the black and dotted line. We also show the MiNNLO$_{\text{PS}}$ prediction in which the Higgs decay is generated by PYTHIA8 (red and dashed curve). This prediction is formally NNLO accurate in the production and only LO accurate in the decay. In the two ratio panels we show the ratio to the MiNNLO$_{\text{PS}}$ central value. In the upper panel, we correlate the scale variation factors when interfacing production and decay events ($K_{R,\mathrm{prod}} = K_{R,\mathrm{dec}}$), so that error bars are obtained through the customary 7-point scale variation. In the lower ratio panel, we do not correlate the scales ($K_{R,\mathrm{prod}} \neq K_{R,\mathrm{dec}}$) and we thus consider all the possible combinations,

---

[1] The corresponding results for $W^-H$ production are shown in Appendix A.

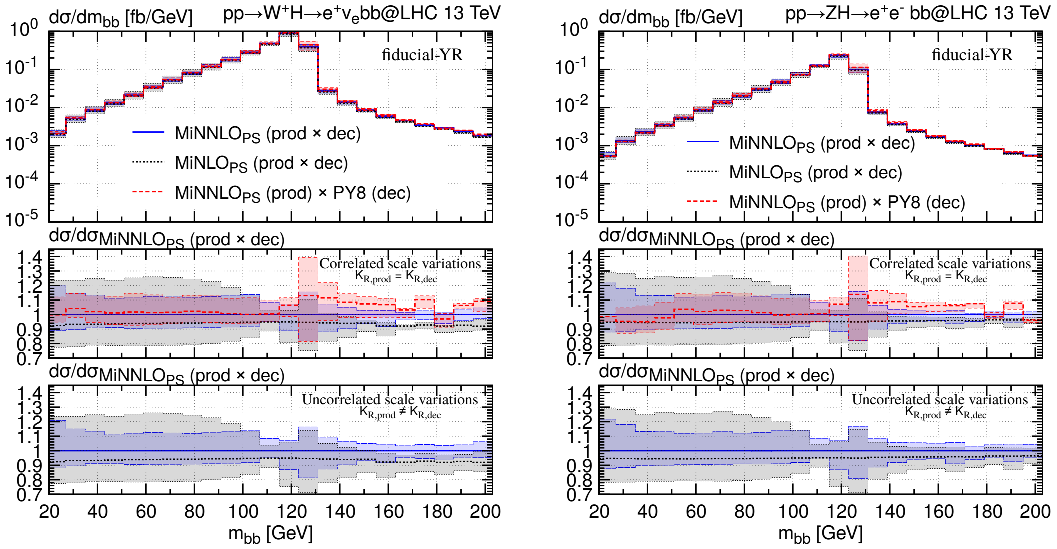


Figure 4.5.: Invariant mass $m_{b\bar{b}}$ of the $b\bar{b}$ system for $e^+\nu_e b\bar{b}$ production (left) and for $e^+e^-b\bar{b}$ production (right) in the `fiducial-YR` setup.

excluding the extreme cases $(0.5, 2)$ and $(2, 0.5)$. This prescription is a more conservative 17-point scale variation. When the shower generates the decay of the Higgs boson (red, dashed curve), the scale variation is handled by PYTHIA8 as well using its so-called *automated parton-shower variation* facility [105]. Since the behaviour observed in $W^{\pm}H$ and $ZH$ is quite similar, we discuss these results simultaneously.

Figure 4.5 shows the invariant mass $m_{b\bar{b}}$ of the $b\bar{b}$ system for $W^+H$ (left) and $ZH$ (right) productions. The MiNNLO$_{\text{PS}}$ and MiNLO$'$ predictions are similar in shape, with a 5% difference in the normalization induced by the MiNNLO$_{\text{PS}}$ corrections. The two results are compatible within scale uncertainties both when correlating and uncorrelating the scale variation factors. Looking at the two ratio panels, we note that using a 7-point or a 17-point scale variation is irrelevant, as error bars are very similar in the two cases. There is an important reduction of scale uncertainties in the MiNNLO$_{\text{PS}}$ results compared to the MiNLO$'$ predictions below the Higgs-mass threshold, while error bars are similar above the threshold. We observe that using PYTHIA8 for simulating the decay of the Higgs boson is a rather good approximation for this observable.

The absolute rapidity difference of the two bottom quarks is shown in figure 4.6. Again, the left panel is $W^+H$ production while the right panel is the $ZH$ case. Firstly, we observe that the MiNNLO$_{\text{PS}}$ and the MiNLO$'$ results are not compatible when we correlate the scales. In fact, we observe a non-reliable error band (roughly $\pm 1\%$) in the MiNNLO$_{\text{PS}}$ result for small values of $|\Delta y_{b,b}|$. Using the more conservative option of not correlating the scales, we obtain more realistic scale uncertainties and we restore the perturbative idea of having the MiNNLO$_{\text{PS}}$ result covered by the MiNLO$'$ error bar. For this observable, the LO decay of the

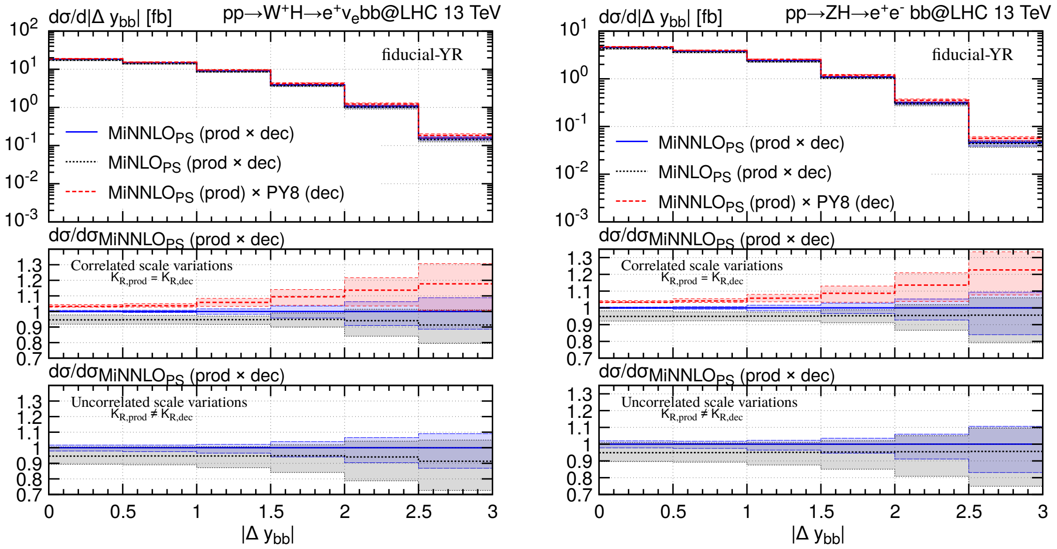


Figure 4.6.: Absolute difference in rapidity $|\Delta y_{b,b}|$ of the $b\bar{b}$ system for $e^+\nu_e b\bar{b}$ production (left) and for $e^+e^-b\bar{b}$ production (right) in the `fiducial-YR` setup.

Higgs boson simulated by PYTHIA8 does not compare well with the MiNNLO$_{\text{PS}}$ prediction, especially in the forward $|\Delta y_{b,b}|$ region.

Similar features can be observed in the other differential observables ($p_{\text{T},bb}$, $p_{\text{T},b_{\text{hard}}}$, $p_{\text{T},e^+}/p_{\text{T},e^-}$ and $m_{WH}/m_{ZH}$) shown in figures 4.7 and 4.8. In all cases we observe a mostly flat $\sim 5\%$ correction induced by MiNNLO$_{\text{PS}}$ on top of the MiNLO$'$ calculation. The MiNNLO$_{\text{PS}}$ predictions have much smaller error bars. In most observables, correlating the scales between production and decay leads to an accidental cancellation that generates unreliable error bars (see e.g. the MiNLO$'$ $p_{T,bb}$ and $p_{\text{T},b_{\text{hard}}}$ distributions around 200 GeV). A more realistic picture is restored when uncorrelating the scales. The inclusion of the Higgs decay using PYTHIA8 produces distributions with shapes similar to the MiNNLO$_{\text{PS}}$ predictions but the normalization is shifted upwards of roughly 5%. This distribution with LO decay is not covered by the MiNNLO$_{\text{PS}}$ error bars in many phase space regions.

### 4.4.4. Comparison to data

In this section we perform a comparison between our MiNNLO$_{\text{PS}}$ predictions and ATLAS data [100] in the in the `fiducial-ATLAS` setup. This comparison is presented in figure 4.9, with the corresponding cross sections shown in table 4.7. For this study we consider all the possible leptonic final states, i.e. $\ell^\pm\nu_\ell b\bar{b}$, $\ell^+\ell^- b\bar{b}$ and $\nu_\ell\bar{\nu}_\ell b\bar{b}$ with $\ell \in \{e, \mu, \tau\}$, and we include hadronization effects through PYTHIA8. However, we note that hadronization effects are small.

We find a good agreement between the MiNNLO$_{\text{PS}}$ results and ATLAS data but we note that experimental measurements have very large uncertainties caused by low statistics, which is due to the large lower cuts on the transverse momentum of the vector boson. We notice that in the original publications of the observation of this process [106, 107], the cut applied was much lower (150 GeV) but no unfolding to the fiducial cross section was performed. Repeating this comparison with data at higher statistics will be desirable in the future.

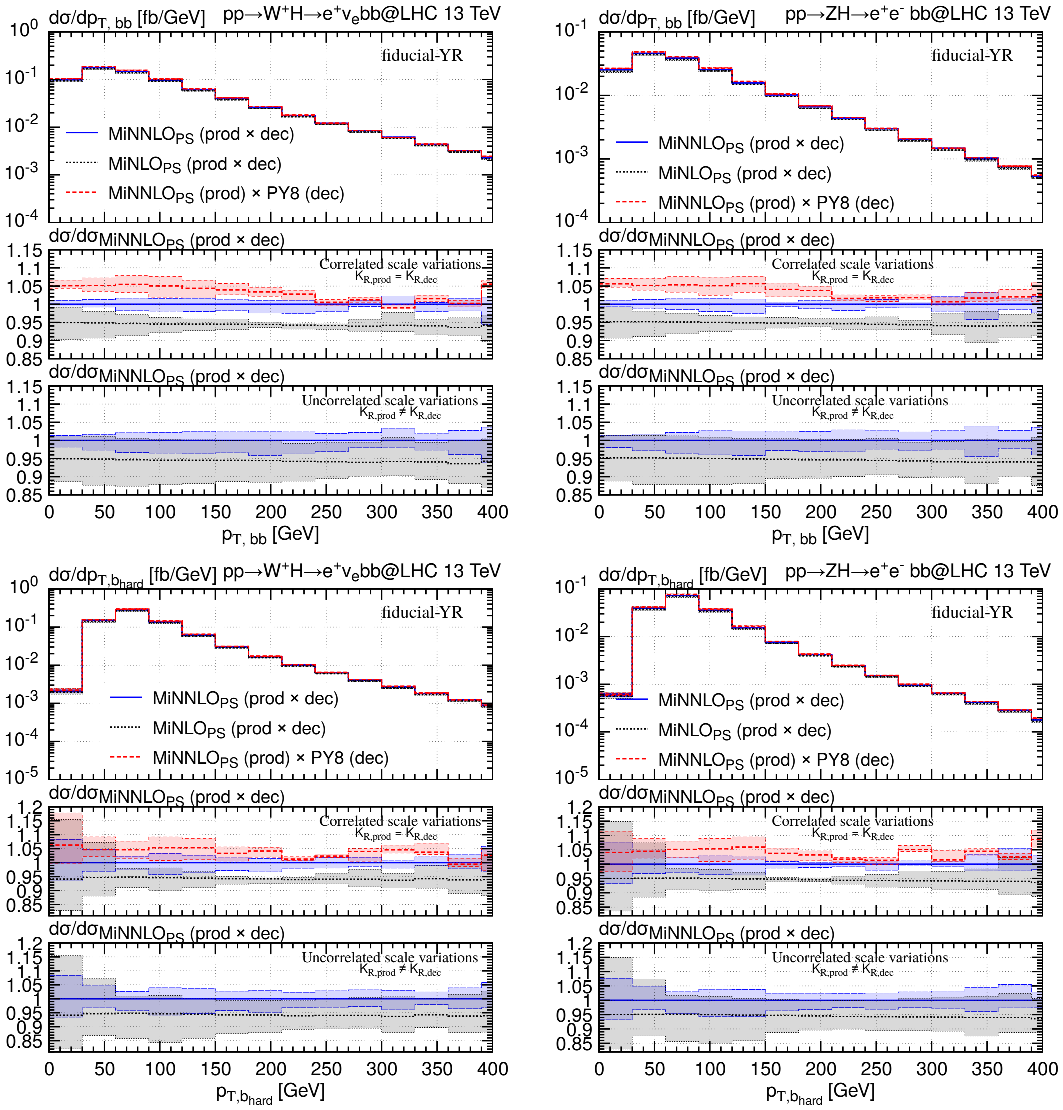


Figure 4.7.: Transverse momentum of the $b\bar{b}$ system $p_{\mathrm{T},bb}$ and of the leading $b$-jet $p_{\mathrm{T},b_{\mathrm{hard}}}$ for $e^+\nu_e b\bar{b}$ (left) and $e^+e^-b\bar{b}$ production (right) in the `fiducial-YR` setup.

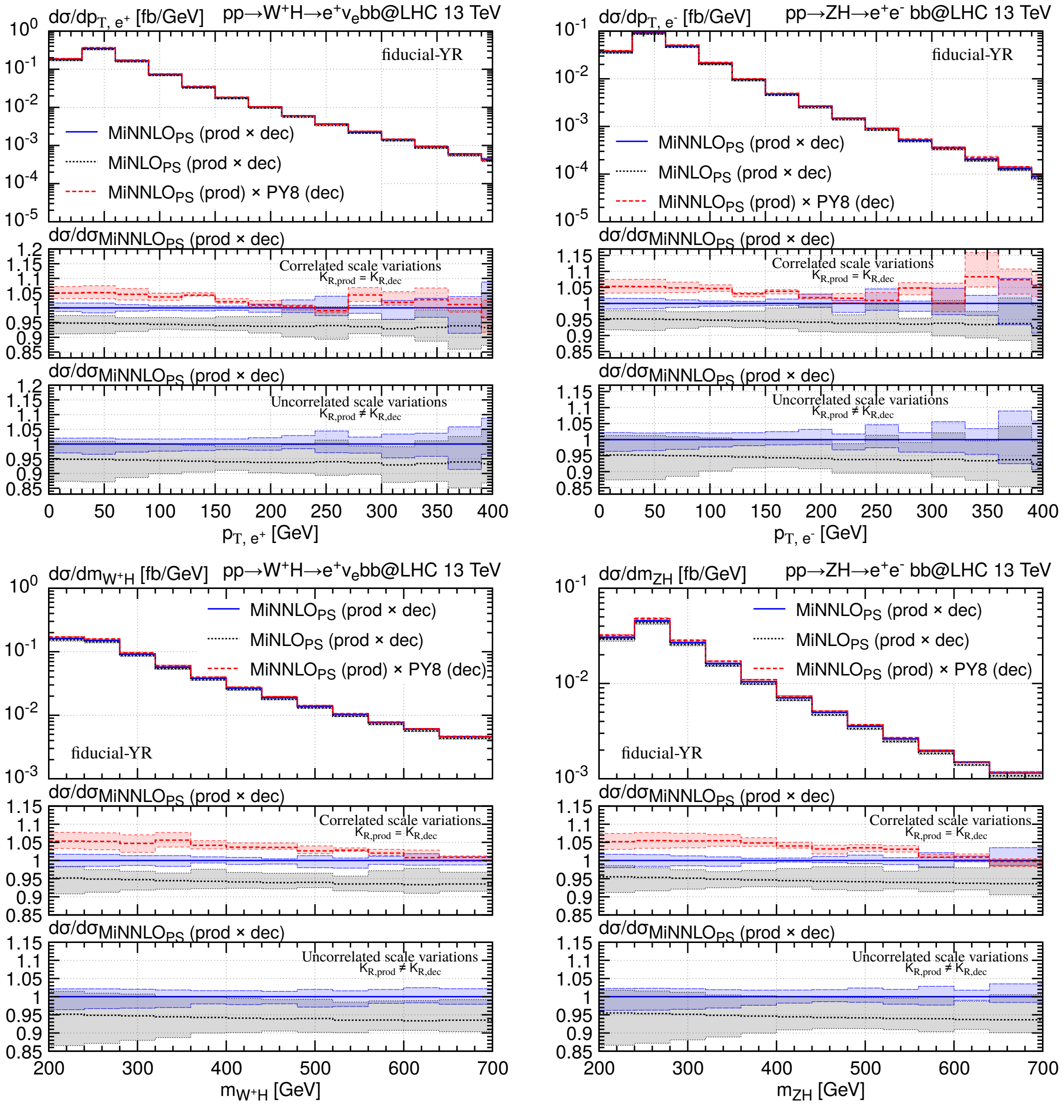


Figure 4.8.: Transverse momentum of the electron/positron $p_{\mathrm{T},e^+}/p_{\mathrm{T},e^-}$ and invariant mass of the colour singlet $m_{WH}/m_{ZH}$ for $e^+\nu_e b\bar{b}$ (left) and for $e^+e^-b\bar{b}$ production (right) in the `fiducial-YR` setup.

| $pp \to W^{\pm}H \to \ell^{\pm}\nu_{\ell}b\bar{b}$ | | |
|---|---|---|
| $\sigma$ [fb] | $p_T^W \in [250{,}400]$ GeV | $p_T^W \in [400, \infty]$ GeV |
| MiNNLO$_{\text{PS}}$ | $6.52^{+2.4\%}_{-1.8\%}$ | $1.46^{+2.5\%}_{-1.9\%}$ |
| ATLAS [100] | $3.3^{+3.6(\text{Stat.})+3.2(\text{Syst.})}_{-3.4(\text{Stat.})-3.0(\text{Syst.})}$ | $2.1^{+1.0(\text{Stat.})+0.6(\text{Syst.})}_{-0.9(\text{Stat.})-0.5(\text{Syst.})}$ |
| $pp \to ZH \to (\ell^+\ell^-, \nu_\ell\bar{\nu}_\ell)b\bar{b}$ | | |
| $\sigma$ [fb] | $p_T^Z \in [250{,}400]$ GeV | $p_T^Z \in [400, \infty]$ GeV |
| MiNNLO$_{\text{PS}}$ | $3.98^{+7.6\%}_{-5.4\%}$ | $0.79^{+6.5\%}_{-4.2\%}$ |
| ATLAS [100] | $1.4^{+2.4(\text{Stat.})+1.9(\text{Syst.})}_{-2.3(\text{Stat.})-1.7(\text{Syst.})}$ | $0.2^{+0.6(\text{Stat.})+0.3(\text{Syst.})}_{-0.5(\text{Stat.})-0.3(\text{Syst.})}$ |

Table 4.7.: Fiducial cross sections in the `fiducial-ATLAS` setup compared to ATLAS data.

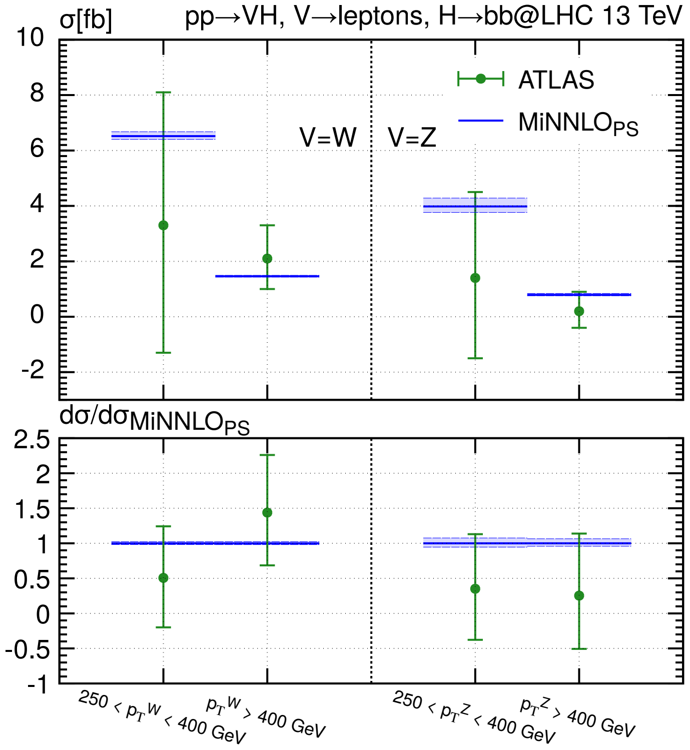


Figure 4.9.: Comparison of MiNNLO$_{\text{PS}}$ predictions to ATLAS data [100] in the `fiducial-ATLAS` setup.

## 4.5. Jet flavour

### 4.5.1. Infrared-safe definition of jet flavour

Identifying the flavour of jets is fundamental for precision phenomenology. If we can correctly assign a flavour to a reconstructed jet, we can significantly improve our ability in selecting specific scattering processes and reject backgrounds. Even though defining a *gluon jet* or a *quark jet* is intuitively clear, providing an infrared-safe definition of the flavour of jets is not trivial. When going beyond LO in QCD, it is necessary to define a procedure that is able to cluster partons into jets and associate each jet with a *well-defined* flavour. Well-defined flavour means a flavour that is insensitive to extra soft and collinear emissions, thus infrared and collinear safe.

In a simple and naive picture, we can define the flavour of jets as follows: firstly, we reconstruct jets using a customary jet-clustering algorithm and we then assign to each jet a net flavour content given by the difference between the total number of quarks and anti-quarks for every possible flavour. If the net content of flavour of a jet is zero, this jet is a gluon jet. If the net content of flavour is (minus) one unit of a possible flavour, the jet is a (anti-) quark jet of that flavour. This procedure is clear and intuitive, but it necessarily leads to an ill-defined definition of flavour of jets. As explained in [99], we can analyse this issue considering the simple process $e^+e^- \to q\bar{q}$. At the first order in the strong coupling, one of the two final-state partons emits a gluon. Applying a usual clustering algorithm (e.g. the $k_T$-clustering algorithm [108]), the gluon can be clustered with the quark, with the anti-quark or it can be considered as a separate jet. Using the naive picture described before for assigning the flavour of jets, one is always able to reconstruct the correct 2→2 Born flavour structure, or to define the process as irreducibly 2→3. At the second order in the strong coupling, this procedure breaks: a large-angle soft gluon can split into a pair of quarks widely separated in angle. These two quarks can be wrongly clustered into two different jets (and not recombined together in a gluon-jet first). This wrong recombination contaminates the flavour of the reconstructed jets. This possible configuration is depicted in figure 4.10 taken from [99]. Given that this problem can appear for arbitrarily soft gluons that split into $q\bar{q}$-pairs, the above definition of jet flavour is IR unsafe from $\mathcal{O}(\alpha_s^2)$ onwards.

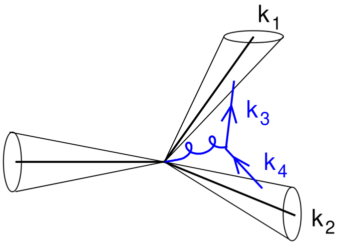


Figure 4.10.: Possible $\mathcal{O}(\alpha_s^2)$ configuration in which a pair of widely separated soft quarks ($k_3$ and $k_4$) coming from a large-angle soft gluon spoils the flavour of jets ($k_1$ and $k_2$). The picture is taken from [99].

Most LHC measurements rely on the anti-$k_T$ algorithm [104] for the reconstruction of jets and the flavour of jets is assigned with a naive procedure (a $b$-jet is a jet with at least one $b$-hadron), thus a reasonable and robust data/theory comparison cannot be performed. In [99], the authors proposed a new clustering algorithm that is IR safe in the assignment of the flavour. This algorithm is called flavour-$k_T$ algorithm and in this section we present its main features.

The starting point of our discussion is the IR and collinear structure of QCD. The usual $k_T$ algorithm tries to reconstruct the QCD splitting history using a distance measure that starts the recombination procedure from particles with the highest probability to have branched. The clustering thus starts from the recombination of particles that are soft or collinear, as this reflects the IR and collinear enhancement of QCD amplitudes. The $k_T$-algorithm does not distinguish between the kind of splitting at hand: a soft quark has the same closeness as a soft gluon, but this does not reflect the QCD behaviour. In fact, the emission of a gluon presents divergences in both soft and collinear regions while the emission of a quark has only the collinear divergence. If we consider $g \to q\bar{q}$, the product of phase space and matrix element can be written as (see section 2.1.3):

$$\mathrm{d}k_j \, |M_{g\to q_i \bar{q}_j}|^2 \simeq \frac{\alpha_s T_R}{2\pi} \frac{\mathrm{d}E_j}{E_i} \frac{\mathrm{d}\theta_{ij}^2}{\theta_{ij}^2} \,, \tag{4.11}$$

with $E_j \ll E_i$ and $\theta_{ij} \ll 1$. We note that the soft divergence is not present. The same conclusion holds also for $q \to qg$ splittings. The $k_T$-algorithm thus introduces a spurious closeness as there is no divergence for $E_j \to 0$. This problem can be solved using a distance measure that depends on the flavour of the particles involved in the clustering procedure.

In the flavour-$k_T$ algorithm, the distance measure $d_{ij}^{(F)}$ between two particles $i$ and $j$ is the following:

$$d_{ij}^{(F)} = \frac{(\Delta\eta_{ij}^2 + \Delta\phi_{ij}^2)}{R^2} \times \begin{cases} \max(p_{\mathrm{T},i}^2, \, p_{\mathrm{T},j}^2)\,, & \text{softer of } i,j \text{ is flavoured,} \\ \min(p_{\mathrm{T},i}^2, \, p_{\mathrm{T},j}^2)\,, & \text{softer of } i,j \text{ is flavourless.} \end{cases} \tag{4.12}$$

The term $(\Delta\eta_{ij}^2 + \Delta\phi_{ij}^2)$ is the usual angular distance between $i$ and $j$ and $R$ is the radius of the cone in the clustering sequence. By *softer* we mean the particle with lower transverse momentum.

To define the distance with the beam, we need to separate explicitly the beam moving towards positive rapidities $B$ (right-moving) and the beam moving toward negative rapidities $\bar{B}$ (left-moving), as we want to define a flavour also for the beam. The distances with $B$ and $\bar{B}$ are thus defined as:

$$\begin{aligned} d_{iB}^{(F)} &= \begin{cases} \max(p_{\mathrm{T},i}^2, \, p_{\mathrm{T},B}^2)\,, & i \text{ is flavoured,} \\ \min(p_{\mathrm{T},i}^2, \, p_{\mathrm{T},B}^2)\,, & i \text{ is flavourless.} \end{cases} \\ d_{i\bar{B}}^{(F)} &= \begin{cases} \max(p_{\mathrm{T},i}^2, \, p_{\mathrm{T},\bar{B}}^2)\,, & i \text{ is flavoured,} \\ \min(p_{\mathrm{T},i}^2, \, p_{\mathrm{T},\bar{B}}^2)\,, & i \text{ is flavourless.} \end{cases} \end{aligned} \tag{4.13}$$

The terms $p_{\mathrm{T},B}$ ($p_{\mathrm{T},\bar{B}}$) represents the hard transverse scale associated with the right-moving (left-moving) beam, defined as follows:

$$p_{\mathrm{T},B}(\eta) = \sum_i p_{\mathrm{T},i} \left( \Theta(\eta_i - \eta) + \Theta(\eta - \eta_i) e^{\eta_i - \eta} \right) , \tag{4.14}$$

$$p_{\mathrm{T},\bar{B}}(\eta) = \sum_i p_{\mathrm{T},i} \left( \Theta(\eta - \eta_i) + \Theta(\eta_i - \eta) e^{\eta - \eta_i} \right) . \tag{4.15}$$

The idea behind equations (4.14) and (4.15) is the following: starting from a DGLAP approach in which dominant contributions are strongly ordered in angle, the beam hardness should be constructed as a function of the rapidity. Moreover, it should be harder than all the emissions that already occurred and dependent on the light-cone momentum still left in the beam.

Given the definitions of the distance between two particles $d_{ij}^{(F)}$ and the distance with the beams $d_{i,B/\bar{B}}^{(F)}$, the clustering procedure is equal to any sequential clustering algorithm: we calculate the distances $d_{ij}^{(F)}$ for every pair of particles and the distances $d_{i,B/\bar{B}}^{(F)}$ of every particle with the beams. If the small calculated distance is a $d_{ij}^{(F)}$, the two particles $i$ and $j$ are clustered together into a particle whose flavour is given by the sum of the two individual flavours. When the smallest distance is $d_{iB}^{(F)}$ or $d_{i\bar{B}}^{(F)}$, the particle is considered a final jet and removed from the clustering sequence.

We can thus apply explicitly this procedure to the problematic configuration depicted in figure 4.10: using the flavour-$k_T$ algorithm, the two widely separated soft quarks are clustered together first, giving rise to a gluon jet. This gluon jet can then be clustered to the other particles involved in the process without spoiling the flavour of jets.

For a complete proof of the IR safety of this algorithm, the interested reader is referred to the original publication [99].

### 4.5.2. Impact of the jet-clustering algorithm

In this section we discuss the impact of using different clustering algorithms for the identification of $b$-jets in the phenomenological analysis presented in section 4.4. In particular, we show a comparison among the anti-$k_T$ algorithm [104] (blue and solid curve), the $k_T$ algorithm [109] (red and short-dashed curve) and the flavour-$k_T$ algorithm [99] (green and long-dashed curve). We remind the reader that the main difference between the anti-$k_T$/$k_T$ algorithms and the flavour-$k_T$ one in the reconstruction of $b$-jets is the fact that in the first case a $b$-jet arises anytime there is at least a bottom (or anti-bottom) quark in a jet, while in the flavour-$k_T$ case there must be an uneven number of $b$ or $\bar{b}$ quarks. In this way, when a $b\bar{b}$-pair is recombined together, we obtain a $b$-jet in the anti-$k_T$/$k_T$ case, while we obtain a gluon-jet with the flavour-$k_T$ algorithm. This different behaviour in the reconstruction of $b$-jets determines large discrepancies in some regions of the phase space.

We consider $W^+H$ production in the `fiducial-YR` setup defined in table 4.5. The radii of the clustering algorithms are always set to $R = 0.4$. We show plots for the invariant mass ($m_{bb}$) and the transverse momentum ($p_{\mathrm{T},bb}$) of the $b\bar{b}$ system, the transverse momentum of the hardest $b$-jet ($p_{\mathrm{T},b_{\mathrm{hard}}}$) and the transverse momentum of the positron ($p_{\mathrm{T},e^+}$).

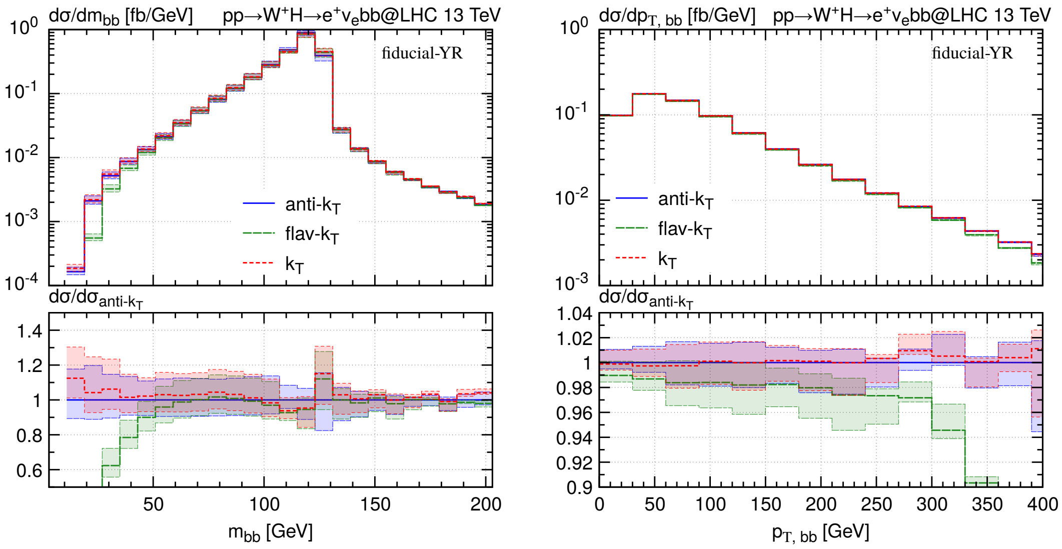


Figure 4.11.: Invariant mass $m_{bb}$ and transverse momentum $p_{\mathrm{T},bb}$ of the $b\bar{b}$ system for $e^+\nu_e b\bar{b}$ production in the `fiducial-YR` setup using the anti-$k_T$, $k_T$ and flavour-$k_T$ algorithms with $R = 0.4$.

Before describing the main differences induced by using a flavour-specific clustering algorithm, we observe that the difference between the anti-$k_T$ and $k_T$ algorithms is small, as expected. For this reason, we will comment only on the difference between the flavour-$k_T$ and $k_T$ algorithms, as they share the same underlying distance in the ordering of the clustering.

The left panel of figure 4.11 shows the invariant mass $m_{bb}$ of the $b\bar{b}$ system. We observe a large discrepancy in the low mass region which is related to configurations in which the two $b$-quarks are separated by a small $\Delta R_{bb}$ distance. In these configurations, the output of the $k_T$ and flavour-$k_T$ clusterings is different, because both algorithms tend to cluster the two bottom quarks together but in the first case we obtain a $b$-jet while in the second one we obtain a gluon-jet. When applying the fiducial cuts of the `fiducial-YR` setup, we require at least two $b$-jets, thus it is more likely to accept the event in the $k_T$ case. Moreover, a soft $b$-quark can be sufficient, as it can be clustered with a hard gluon, giving rise to a sufficiently hard $b$-jet that passes the fiducial cuts. This recombination is instead strongly suppressed when applying the flavour-$k_T$ clustering. Furthermore, the two algorithms behave differently even when $\Delta R_{bb} > R$: the $k_T$ algorithm tends to cluster two different $b$-jets, while the flavour-$k_T$ algorithm can cluster the two $b$-quarks together because of the definition of the beam distance. In particular, the algorithm compares the relative distance $d_{ij}^{(F)}$ with the distance of each bottom quark to the beams $d_{bB}^{(F)}$ and $d_{\bar{b}\bar{B}}^{(F)}$. By construction, the distance with the beams is greater than the transverse momenta of the bottom (anti-bottom) quarks, so $d_{ij}^{(F)}$ can be smaller than $d_{bB}^{(F)}$ and $d_{\bar{b}\bar{B}}^{(F)}$), even for $\Delta R_{bb} > R$.

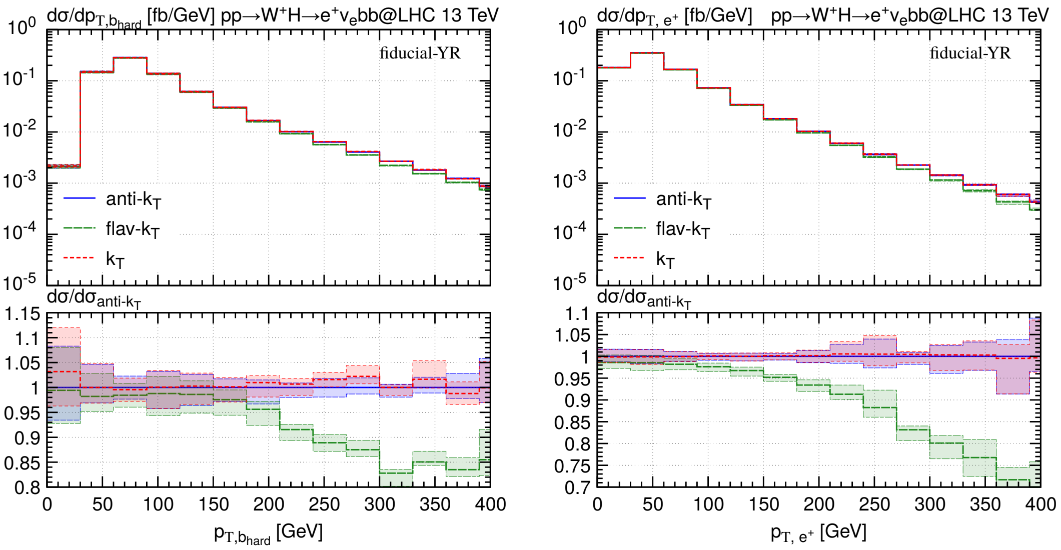


Figure 4.12.: Transverse momentum of the leading $b$-jet $p_{\mathrm{T},b_{\mathrm{hard}}}$ and of the positron $p_{\mathrm{T},e^+}$ for $e^+\nu_e b\bar{b}$ production in the `fiducial-YR` setup using the anti-$k_T$, $k_T$ and flavour-$k_T$ algorithms with $R = 0.4$.

The same motivations explain the large discrepancies in the $p_{\mathrm{T},bb}$ distribution (right panel in figure 4.11) and the $p_{\mathrm{T},b_{\mathrm{hard}}}$ distribution (left panel in figure 4.12). In fact, there are large differences in the high-$p_{\mathrm{T}}$ regions, where the Higgs boson is highly boosted and thus tends to decay into a collimated pair of $b$-quarks or an energetic leading $b$-quark.

The right panel of figure 4.12 shows that the clustering algorithm for the reconstruction of $b$-jets can also have an impact on observables that are not directly associated with bottom quarks. In fact, we show that we observe a large discrepancy in the $p_{\mathrm{T},e^+}$ distribution for large values of the transverse momentum. These differences are again associated to configurations with a highly boosted Higgs, as the positron recoils on the Higgs system and becomes boosted as well.

We conclude by stating that our plots clearly show that the impact of the clustering algorithm that is used for the identification of $b$-jets is absolutely not negligible as it introduces large discrepancies when compared to usual jet-clustering algorithms. We thus suggest reconstructing heavy-quark jets through the flavour-$k_T$ algorithm only when it is used also on the experimental side, which is usually not the case because of its complexity (see [110] for a detailed analysis on the applicability of this algorithm). A possible (but not always feasible) solution is given by performing a massive calculation in which the kinematics of the $b$-quarks is described retaining a full dependence on the mass at the amplitude level. This approach is commonly referred to as *four-flavour scheme* (4FS). The main drawback is that massive calculations are not always available at high accuracy. A matching between precise massless computations and parton showers thus becomes fundamental, as we can perform a reshuffling of momenta at the parton shower level in order to restore the correct masses. However, this reshuffling is not completely unambiguous and it introduces uncertainties on the kinematics of the bottom quarks.

# 5. $ZH$ production with $H \to b\bar{b}$ decay at NNLO+PS accuracy in the SMEFT

## 5.1. Motivation

The current formulation of the Standard Model cannot be the final fundamental theory of nature. Even though it appears to be incredibly precise and predictive in the description of scattering processes at colliders, many experimental and theoretical issues cannot be explained through the SM. For example, in the SM neutrinos are massless particles, which is in contradiction with flavour oscillations observed experimentally [111, 112]. The SM cannot explain the presence of Dark Matter and Dark Energy that are detected through cosmological observations [113]. Furthermore, the abundance of matter over antimatter in the Universe suggests the presence of additional sources of charge-conjugation and parity (CP) violation other than the SM mechanism [114]. Moreover, according to our current knowledge, the SM cannot be unified with the theory of General Relativity.

Despite its incompleteness, the striking success of the SM at the LHC led to the possibility of formulating a final fundamental theory of the subatomic world through an effective-field-theory (EFT) in which the SM is the low-energy manifestation of a complete UV theory beyond the SM (BSM). In this framework, the SM physics at the electroweak scale is *decoupled* from the complete UV theory, which means that the two dynamics are factorized because they occur at largely different scales. The physical motivation is that two phenomena widely separated in energy do not affect each other, up to some corrections proportional to the ratio of the two involved scales [115].

An EFT framework reveals to be extremely useful when the complete UV theory is not known. In fact, it is possible to use a *bottom-up* approach to parametrize our lack of knowledge of the complete theory, without constructing a specific BSM model. In this way, we simply parametrize deviations from the SM predictions that can be matched to the complete theory afterwards. A widely used BSM theory constructed using this logic is given by the Standard Model Effective Field Theory (SMEFT) [116, 117, 118].

In the SMEFT, we correct the low-energy dynamics as predicted by the SM with a set of local operators that respect the same $\mathrm{SU}(3)_\mathrm{C} \times \mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_\mathrm{Y}$ gauge symmetry structure. The SMEFT Lagrangian thus reads

$$\mathcal{L}_{\mathrm{SMEFT}} = \mathcal{L}_{\mathrm{SM}} + \sum_{n=5}^{\infty} \sum_{i} \frac{C_i^{(n)}}{\Lambda^{n-4}} \mathcal{O}_i^{(n)} \,, \tag{5.1}$$

where $\mathcal{O}_i^{(n)}$ are local operators with mass dimension $n$, $C_i^{(n)}$ are the corresponding unknown Wilson coefficients and $\Lambda$ represents the new physics scale. The Wilson coefficients are unspecified and constrained by experiments: a Wilson coefficient different from zero is an indication of new physics. Once bounds on the Wilson coefficients are provided, we can obtain bounds on the parameters of the BSM theory.

BSM searches constitute an important part of the current and future LHC physics programme. As already presented in chapter 4, the Higgs boson is crucial in this context because a better understanding of the Higgs sector would lead to an improvement of our knowledge of the theory and a possible extension towards a BSM model. In the SM, the largest branching fraction of the Higgs boson is the $H \to b\bar{b}$ decay ($\sim 60\%$) and the highest sensitivity to this decay is given by the production of the Higgs boson through Higgsstrahlung (see section 4.1). This process has been observed by both the ATLAS and CMS collaborations [106, 107]: the $H \to b\bar{b}$ signal strength in the Higgstrahlung channel ( $\mu^{VH}_{b\bar{b}}$) is constrained to be as predicted by the SM within 25% at one standard deviation. When Run III data will be available and, even more so, during the high-luminosity phase, this signal strength will become much more constraining: the ultimate projected accuracy for $\mu^{WH}_{b\bar{b}}$ and $\mu^{ZH}_{b\bar{b}}$ is 15% and 5%, respectively. The $pp \to ZH \to \ell^+\ell^- b\bar{b}$ process is thus particularly interesting for BSM searches in the Higgs sector.

In this chapter, we present an extension of the calculation performed in chapter 4 for $pp \to ZH \to \ell^+\ell^- b\bar{b}$ in the SMEFT. Both $pp \to ZH$ and $H \to f\bar{f}$ (with $f$ generic fermion) have been studied in the SMEFT at NLO QCD and NLO EW in different works (see e.g. [119, 120, 121, 122, 123, 124]). In the framework of anomalous couplings, the full process $pp \to ZH \to \ell^+\ell^- b\bar{b}$ has been studied at NNLO QCD in [125]. In this chapter, we show how to obtain NNLO+PS predictions in the SMEFT for $pp \to ZH \to \ell^+\ell^- b\bar{b}$ including a set of dimension-six operators that have a direct impact in QCD, without considering EW effects. The results described in this chapter were originally presented by the author and collaborators in [3].

This chapter is organised as follows: we first describe the calculation in section 5.2 and we present in detail our practical implementation in section 5.3. In 5.4 we present a phenomenological analysis for LHC collisions, discussing which kinematic observables are particularly suitable to enhance our sensitivity to BSM effects related to the operators considered in the calculation.

## 5.2. Outline of the calculation

We consider $ZH$ production

$$pp \to ZH \to \ell^+\ell^- H\,, \tag{5.2}$$

with the subsequent decay of the Higgs boson $H$ to a pair of bottom quarks:

$$H \to b\bar{b}\,. \tag{5.3}$$

With $\ell$ we indicate a massless lepton ($\ell \in \{e, \mu, \tau\}$). We consider the following SMEFT operators that modify the SM Higgs field and couplings up to NNLO QCD accuracy:

$$\begin{aligned} Q_{H\Box} &= (H^\dagger H)\,\Box\,(H^\dagger H)\,, & Q_{HD} &= (H^\dagger D_\mu H)^*\,(H^\dagger D^\mu H)\,, \\ Q_{bH} &= y_b\,(H^\dagger H)\,\bar{q}_L\,b_R\,H\,, & Q_{bG} &= \frac{g_s^3}{(4\pi)^2}\,y_b\,\bar{q}_L\sigma_{\mu\nu}T^a\,b_R\,H G^{a,\mu\nu}\,, \\ Q_{HG} &= \frac{g_s^2}{(4\pi)^2}\,(H^\dagger H)\,G^a_{\mu\nu}\,G^{a,\mu\nu}\,, & Q_{3G} &= \frac{g_s^3}{(4\pi)^2}\,f^{abc}\,G^{a,\nu}_\mu G^{b,\sigma}_\nu G^{c,\mu}_\sigma\,. \end{aligned} \tag{5.4}$$

They enter the full SMEFT Lagrangian as:

$$\mathcal{L}_{\text{SMEFT}} \supset \sum_i \frac{C_i}{\Lambda^2}\,Q_i\,. \tag{5.5}$$

These operators $Q_i$ have dimension six and their product with the corresponding Wilson coefficients $C_i$ is suppressed by two powers of the new physics scale $\Lambda$. Notice that we do not include any EW operators (e.g. operators that modify the coupling of the Higgs boson with two $Z$ bosons) and $Q_{3G}$ does not directly affect the Higgs boson but it is needed to reach NNLO QCD accuracy.

With $H$ we denote the SM Higgs doublet, $q_L$ is the left-handed third-generation quark doublet and $b_R$ is the right-handed bottom-quark singlet. The QCD coupling constant is $g_s = \sqrt{4\pi\alpha_s}$. $G^a_{\mu\nu}$ represents the QCD field strength tensor, where $a$ is a colour index and $\mu, \nu$ are Lorentz indices. Given the SU(3) generators $T^a$, the covariant derivative reads: $D_\mu = \partial_\mu - i\,g_s\,G^a_\mu\,T^a$. With $f^{abc}$ we denote the QCD antisymmetric structure constant. We also use $\Box = \partial_\mu\,\partial^\mu$ and $\sigma_{\mu\nu} = i/2\,(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu)$, with $\gamma_\mu$ the usual Dirac matrices. The bottom-quark Yukawa coupling is $y_b = \sqrt{2}\,\overline{m}_b/v$, where $\overline{m}_b$ is the mass of the bottom quark in the $\overline{\text{MS}}$ scheme and $v$ is the Higgs vacuum expectation value (VEV). In the case of $Q_{bH}$ and $Q_{bG}$ operators, we assume the sum over the hermitian conjugate.

The normalisation adopted in equation (5.4) reveals to be extremely useful. In particular, we explicitly include a factor $y_b$ in both $Q_{bH}$ and $Q_{bG}$, as this factor appears in many UV complete theories that match our operators [126, 127]. Moreover, we include some powers of the strong coupling $g_s$ and a factor $1/(4\pi)^2$ in $Q_{bG}$, $Q_{HG}$ and $Q_{3G}$, in order to have the Wilson coefficients $C_{bG}$, $C_{HG}$ and $C_{3G}$ of order $\mathcal{O}(1)$ in all weakly-coupled UV-complete extensions of the SM. We stress that this normalisation is not the only one possible (see e.g. [122]) but it turns out to be particularly advantageous for a clear power counting in the SMEFT.

Our predictions are obtained following the strategy presented in chapter 4. More precisely, we separately generate production events ($pp \to ZH$) and decay events ($H \to b\bar{b}$) in the SMEFT. The event generation for the production stage consists in the computation of SMEFT corrections that are factorized from the SM calculation; the generation of decay events requires the computation of both factorizable and non-factorizable corrections. We then combine production and decay events using a narrow-width approximation formula, as explained in 4.2.3. Once the complete events are obtained, we shower them using PYTHIA8 [87] applying a dedicated veto procedure. In our computation, bottom quarks are considered massless, but we keep a non-zero bottom Yukawa coupling. We impose the minimal-flavour violation hypothesis [128] and set the Cabibbo-Kobayashi-Maskawa matrix element $V_{tb}$ to one.

### 5.2.1. SMEFT factorizable contributions to $H \to b\bar{b}$ decay

The SMEFT operators $Q_{H\Box}$, $Q_{HD}$ and $Q_{bH}$ are associated to factorizable contributions: the corresponding matrix elements can be obtained from the massless $H \to b\bar{b}$ computation in the SM at NNLO accuracy [83, 84, 129] with a rescaling of the bottom Yukawa coupling $y_b$

$$y_b^2 \to y_b^2 \left(1 + 2\, c_{\text{fac}}\right), \tag{5.6}$$

where

$$c_{\text{fac}} = c_{\text{kin}} - c_{bH}, \qquad c_{\text{kin}} = \frac{v^2}{\Lambda^2}\left[C_{H\Box} + \frac{C_{HD}}{4}\right], \qquad c_{bH} = \frac{v^2}{\Lambda^2}\,\text{Re}\left(C_{bH}\right). \tag{5.7}$$

The effect of $Q_{H\Box}$, $Q_{HD}$ and $Q_{bH}$ factorizes to all orders in the strong coupling because these operators do not contain gluons. In equation (5.7), the term $c_{\text{kin}}$ is associated to the canonical normalisation of the Higgs kinetic term when $Q_{H\Box}$ and $Q_{HD}$ are included in the calculation. The Higgs VEV $v$ is defined as

$$v^2 = 1/\left(\sqrt{2} G_F\right), \tag{5.8}$$

and it is thus related to the Fermi constant $G_F$ which is obtained from muon decay. The relation between the VEV and $G_F$ is modified at LO by EW corrections in the SMEFT. These corrections are very small [118, 121] and thus negligible in our study.

Given the simplicity of the replacement in (5.6), we can obtain the analytic $H \to b\bar{b}$ decay rate associated to the inclusion of the operators $Q_{H\Box}$, $Q_{HD}$ and $Q_{bH}$ in a straightforward way, starting from the corresponding results in the SM. At LO accuracy, the SM decay width reads

$$\Gamma(H \to b\bar{b})^{\text{LO}}_{\text{SM}} = \frac{3\, y_b^2\, m_H}{16\pi}, \tag{5.9}$$

and the corresponding NNLO result is [130, 131, 132]:

$$\Gamma(H \to b\bar{b})^{\text{NNLO}}_{\text{SM}} = \left(1 + \Delta\right) \Gamma(H \to b\bar{b})^{\text{LO}}_{\text{SM}}, \qquad \Delta = \frac{\alpha_s}{\pi}\, 5.67 + \left(\frac{\alpha_s}{\pi}\right)^2 29.15. \tag{5.10}$$

Both the $b$-quark Yukawa coupling and the strong coupling constant $\alpha_s$ are evaluated at a scale equal to the Higgs-boson mass $m_H$. Moreover, in the evaluation of the numerical coefficients in the expansion of $\Delta$ in equation (5.10) we consider five active flavours. The SMEFT NNLO result is easily obtained by performing the rescaling in equation (5.6):

$$\Gamma(H \to b\bar{b})^{\text{NNLO,fac}}_{\text{SMEFT}} = (1 + 2\, c_{\text{fac}})\, \Gamma(H \to b\bar{b})^{\text{NNLO}}_{\text{SM}}. \tag{5.11}$$

We note that the inclusive $H \to b\bar{b}$ decay rate in the SM is known up to $\mathcal{O}(\alpha_s^4)$ in the limit of massless $b$-quarks [133, 134, 135], so it is possible to extend equation (5.11) to this order in a straightforward way. Since our goal is to reach NNLO+PS accuracy, equation (5.11) is sufficient for our purpose.

The practical implementation of our computation relies on the MC code used in [61, 2] for reaching NNLO+PS accuracy for the $H \to b\bar{b}$ decay in the SM. The SM matrix elements needed for the computation are taken from [84]. The generation of decay events associated to factorizable corrections is done simply by taking the SM results as obtained in chapter 4 using the NNLOPS method and applying the correct rescaling factor (equation (5.6)) a posteriori, as we will describe in section 5.3.

### 5.2.2. SMEFT non-factorizable contributions to $H \to b\bar{b}$ decay

The insertion of the $Q_{bG}$, $Q_{HG}$ and $Q_{3G}$ operators gives rise to non-factorizable corrections to the fully differential $H \to b\bar{b}$ decay rate.

The operator $Q_{bG}$ is associated to Feynman diagrams of the type shown in figure 5.1. In particular, we show tree-level and one-loop contributions to the $H \to b\bar{b}g$ decay in the upper left and right panels, respectively. Tree-level diagrams in the $H \to b\bar{b}q\bar{q}$ and $H \to b\bar{b}gg$ channels are shown in the central panels. Moreover, the lower diagrams represent two-loop contributions to the $H \to b\bar{b}$ decay. The corresponding matrix elements are obtained interfering the SMEFT diagrams with the relevant SM amplitudes and their explicit expressions can be found in Appendix A of the original publication [3].

The dominant contribution is given by the interference of the tree-level $H \to b\bar{b}g$ correction in SMEFT (upper left diagram in figure 5.1) with the real $H \to b\bar{b}g$ SM amplitude. This $H \to b\bar{b}g$ correction is IR finite. In fact, the corresponding one-loop correction is exactly zero because it involves only scaleless integrals. Using the normalisation defined in equation (5.4), this contribution appears in our calculation at NNLO (more precisely, it is of order $\mathcal{O}(y_b^2\, \alpha_s^2\, C_{bG})$). After integrating over the three-body phase space, we find the following inclusive $H \to b\bar{b}$ decay width

$$\Gamma(H \to b\bar{b})^{\text{NNLO,non}}_{\text{SMEFT}} = \Delta_{\text{non}}\, c_{bG}\, \Gamma(H \to b\bar{b})^{\text{LO}}_{\text{SM}}\,, \qquad \Delta_{\text{non}} = \left(\frac{\alpha_s}{\pi}\right)^2 \frac{m_H^2}{3v^2}\,, \tag{5.12}$$

where we have defined $c_{bG}$ as:

$$c_{bG} = \frac{v^2}{\Lambda^2}\, \text{Re}\left(C_{bG}\right)\,. \tag{5.13}$$

We emphasise that our result agrees with the one presented in [122] after taking the limit for massless bottom quarks.

Using current experimental bounds, the real part of the Wilson coefficient $C_{bG}$ appearing in (5.13) is weakly constrained [136, 137, 138]. In fact, values of $c_{bG} = \mathcal{O}(100)$ are allowed by current data and thus the NNLO correction in equation (5.12) is numerically sizeable, of the order of a usual NLO computation. For this reason, we included in our computations also N$^3$LO corrections ($\mathcal{O}(y_b^2\, \alpha_s^3\, C_{bG})$). At this order, the contributing Feynman diagrams are virtual corrections (upper right panel in 5.1) and real corrections leading to both $H \to b\bar{b}q\bar{q}$ and $H \to b\bar{b}gg$ (central panels in 5.1). Moreover, at this order the two-loop contributions represented by the lower diagrams have to be accounted for as well.

The operator $Q_{HG}$ can be discarded from our computation as it does not contribute to the fully differential decay rate $H \to b\bar{b}$ in the massless limit for the bottom quarks. This is related to the fact that an insertion for this operator does not produce amplitudes with mixed chiralities for the final-state bottom quarks in case of massless quarks. More details can be found in appendix B of [3].

As for the operator $Q_{3G}$, we have both tree-level corrections to $H \to b\bar{b}gg$ and one-loop corrections to $H \to b\bar{b}g$. These contributions are of order $\mathcal{O}(y_b^2\, \alpha_s^3\, C_{3G})$ (i.e. N$^3$LO), so beyond the accuracy of our calculation. Moreover, given the actual bounds on the Wilson coefficient $C_{3G}$ [139], they have a very minor numerical impact (see appendix B of [3] for

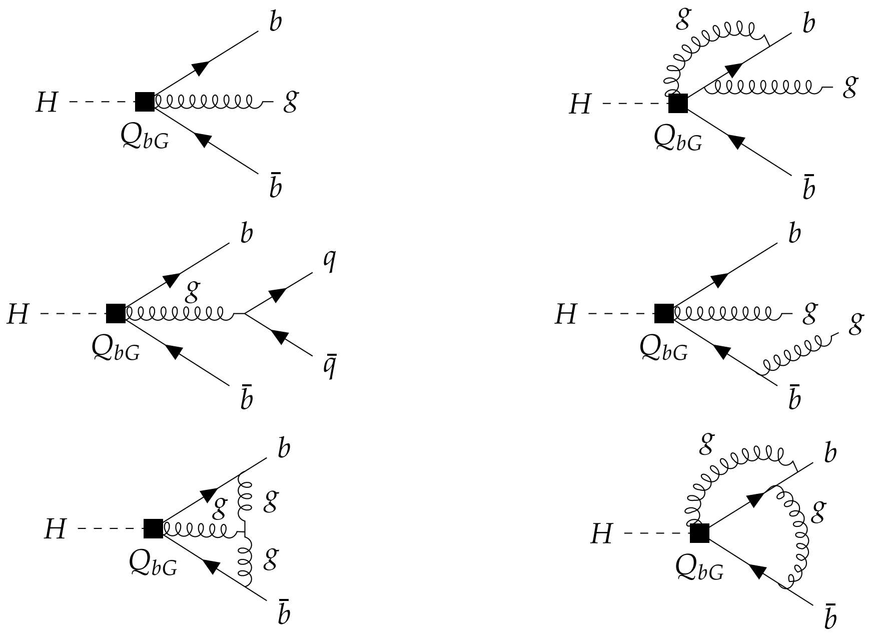


Figure 5.1.: Sample Feynman diagrams contributing to the $H \to b\bar{b}$ decay up to N$^3$LO accuracy with the insertion of the operator $Q_{bG}$.

more details). For these reasons, contributions associated to this operator have been neglected in our computation.

### 5.2.3. SMEFT contributions to $pp \to ZH$

Among the operators defined in equation (5.4), the largest correction to $pp \to ZH$ is given by $Q_{H\Box}$ and $Q_{HD}$. These operators induce a contribution at LO that can be obtained with the following shift:

$$g_{HZZ}^2 \to g_{HZZ}^2 \left(1 + 2\, c_{\text{kin}}\right) . \tag{5.14}$$

With $g_{HZZ}$ we denote the coupling of the Higgs boson with two $Z$ bosons in the SM ($g_{HZZ} = 2\, m_Z^2/v$). We recall that we consider only operators that have an impact up to NNLO QCD accuracy, so that effective EW operators that modify the $g_{HZZ}$ coupling in a non-trivial way are not accounted for in our computation.

Using the shift in equation (5.14), the inclusive cross section for $ZH$ production in the SMEFT reads

$$\sigma(pp \to ZH)_{\text{SMEFT}}^{\text{NNLO}} = (1 + 2\, c_{\text{kin}})\, \sigma(pp \to ZH)_{\text{SM}}^{\text{NNLO}} . \tag{5.15}$$

The contributions coming from the insertion of $Q_{bH}$ and $Q_{bG}$ are exactly zero in the case of massless bottom quarks. The operator $Q_{HG}$ contributes at NNLO with corrections that are lower than the permille level (see appendix B of [3]), so that they do not have a concrete impact on our results and we thus neglected them. Finally, the operator $Q_{3G}$ contributes at N$^3$LO accuracy, so we discarded it in the calculation.

## 5.3. NNLO+PS implementation in the SMEFT

In this section we discuss the practical implementation of a NNLO+PS MC code for obtaining predictions of $pp \to ZH \to \ell^+\ell^- b\bar{b}$ in the SMEFT. The strategy we adopt is similar to what we used for the generation of the SM results for the same process described in chapter 4. First, we separately generate results for the production and the decay stages of the Higgs boson. In this case, for the decay we generate two sets of events: the first one contains the factorizable corrections while the second one is related to non-factorizable contributions. We then combine production and decay events employing a narrow-width approximation for the Higgs boson. Once the events for the complete process are produced, we shower them using PYTHIA8. The application of the shower requires an appropriate veto procedure.

### 5.3.1. Generation of events

The first step for obtaining results for $pp \to ZH \to \ell^+\ell^- b\bar{b}$ production in the SMEFT is the generation of both production and decay events.

- Generation of production events: SMEFT corrections to $pp \to ZH$ production factorize according to equation (5.14). We thus generate SM results through the MiNNLO$_{\mathrm{PS}}$ implementation presented in [2] within the POWHEG-BOX-RES framework [88], following what is done in chapter 4. The correct SMEFT rescaling factor will be included in a later step (more details in the combination section 5.3.2).

- Generation of decay events: the $H \to b\bar{b}$ decay is affected by both factorizable and non-factorizable contributions. The generation of events associated to factorizable corrections is straightforward: we use the SM implementation presented in [2] within the POWHEG-BOX-V2 framework [41], as described in chapter 4. This calculation makes use of the MiNLO$'$ + reweighting procedure. The reweighting factor we adopt is described in equation (4.4). The correct SMEFT rescaling factor (equation (5.6)) is included a posteriori when interfacing production and decay events.

  The generation of events associated to non-factorizable corrections is more involved. We recall that the leading contribution is related to the insertion of the $Q_{bG}$ operator, which contributes at order $\mathcal{O}(\alpha_s^2)$. This contribution is finite when the gluon is unresolved and we can thus perform our computation in the POWHEG framework without the need of the MiNLO$'$ method, at variance with the factorizable corrections. From now on, we will denote this correction as R. We include corrections up to $\mathcal{O}(\alpha_s^3)$ through a usual POWHEG computation: we provide both real and one-loop virtual matrix elements (which we denote RV and RR) as depicted in figure 5.1 (upper right and central panels). IR singularities are handled through the FKS subtraction method, as default in POWHEG. The last missing contribution is given by the two-loop corrections depicted in the lower panels of 5.1, which we call VV. This contribution is not associated to a correction to the $H \to b\bar{b}g$ process, but rather to $H \to b\bar{b}$. Given that it is IR finite, we can integrate over the 2-body phase space, obtaining

$$\Gamma(H \to b\bar{b})^{\text{non,VV}}_{\text{SMEFT}} = \frac{\alpha_s}{\pi}\frac{49}{12}\,\Delta_{\text{non}}\,c_{bG}\,\Gamma(H \to b\bar{b})^{\text{LO}}_{\text{SM}}\,. \tag{5.16}$$

and add this result to the SMEFT factorizable contributions to $H \to b\bar{b}$ in a later step (more details in the combination section). Note that the non-factorizable corrections have a linear dependence on $c_{bG}$. For producing our results, we set $c_{bG} = 1$[1] for the generation of events and we modify this value only in a later step.

### 5.3.2. Combination of production and decay events

The combination of production and decay events is done following the strategy presented in section 4.2.3. Note that events for the production stage are combined with decay events associated to both factorizable and non-factorizable corrections, thus obtaining two sets of complete events: the first set contains in the decay the factorizable corrections, while the second set contains the non-factorizable corrections.

We employ a narrow-width approximation formula for the Higgs boson, so that the weight for the complete event $w^{\text{SMEFT}}_{\text{full}}$ can be obtained in a straightforward way starting from equation (4.7):

$$w^{\text{SMEFT}}_{\text{full}} = \frac{w^{\text{SMEFT}}_{\text{prod}}\,w^{\text{SMEFT}}_{\text{dec}}}{\Gamma^{\text{SMEFT}}_{H}}\,, \tag{5.17}$$

where $w^{\text{SMEFT}}_{\text{prod}}$ is the weight associated to $pp \to ZH$ events and $w^{\text{SMEFT}}_{\text{dec}}$ is the weight for $H \to b\bar{b}$ events. The latter can be either associated to the factorizable or non-factorizable contributions.

In our concrete implementation, we use $\Gamma^{\text{SM}}_{H}$ instead of $\Gamma^{\text{SMEFT}}_{H}$ in the previous equation

$$w^{\text{SMEFT}}_{\text{full}} = \frac{w^{\text{SMEFT}}_{\text{prod}}\,w^{\text{SMEFT}}_{\text{dec}}}{\Gamma^{\text{SM}}_{H}}\,, \tag{5.18}$$

and we then apply the correct rescaling factor $\Gamma^{\text{SM}}_{H}/\Gamma^{\text{SMEFT}}_{H}$ at a later stage. This allows us to modify the Wilson coefficients involved in the calculation a posteriori, without recalculating any of the individual cross sections.

The weights associated to the production events and to the factorizable decay events are the SM weights because the correct SMEFT rescaling factors are accounted for in the next step of the combination procedure. This means that we concretely use $w^{\text{SM}}_{\text{prod}}$ and $w^{\text{SM}}_{\text{dec}}$ in (5.17). As for the decay events with non-factorizable corrections, we remind the reader that $c_{bG}$ has been set to 1 in the event generation: this allows us to modify this Wilson coefficient in a flexible way at the level of differential distributions in a later stage.

Once the two sets of complete events are obtained, we shower them using PYTHIA8 and applying the veto procedure described in section 4.2.3.

---

[1] In our code, $c_{bG}$ is expressed in units of $v^2$.

### 5.3.3. NNLO+PS SMEFT differential cross section

The differential distributions obtained after showering the events can be combined as follows:

$$\begin{aligned} \mathrm{d}\sigma_{\text{NNLO+PS}} = \left(1+2c_{\text{kin}}\right)^2 \Bigg\{ \left[1 - 2c_{bH} + \frac{\Gamma(H \to b\bar{b})^{\text{non,VV}}_{\text{SMEFT}}}{\Gamma(H \to b\bar{b})^{\text{NNLO}}_{\text{SM}}}\right] \mathrm{d}\sigma^{\text{SM}}_{\text{NNLO+PS}} \\ + \mathrm{d}\sigma^{\text{non,R+RV+RR}}_{\text{NNLO+PS}} \Bigg\} \frac{\Gamma^{\text{SM}}_{H}}{\Gamma^{\text{SMEFT}}_{H}} . \end{aligned} \tag{5.19}$$

$\mathrm{d}\sigma^{\text{SM}}_{\text{NNLO+PS}}$ represents the differential cross section obtained interfacing the SM production events with the SM decay events while $\mathrm{d}\sigma^{\text{non,R+RV+RR}}_{\text{NNLO+PS}}$ is the differential cross section obtained interfacing the SM production events with the decay events associated to the non-factorizable R, RV, RR corrections.

The term $\left(1+2c_{\text{kin}}\right)^2$ is the SMEFT rescaling that accounts for the correct normalisation of the Higgs kinetic term. It comes from the production (5.15) and factorizable contributions in the decay (5.11) stages.

We stress that both $\Gamma(H \to b\bar{b})^{\text{non,VV}}_{\text{SMEFT}}$ and $\mathrm{d}\sigma^{\text{non,R+RV+RR}}_{\text{NNLO+PS}}$ depend linearly on $c_{bG}$. Note that we now include the correct width factor $\Gamma^{\text{SM}}_{H}/\Gamma^{\text{SMEFT}}_{H}$. The total SMEFT decay width $\Gamma^{\text{SMEFT}}_{H}$ is given by

$$\begin{aligned} \Gamma^{\text{SMEFT}}_{H} = \left(1+2c_{\text{kin}}\right) \Big[ \Gamma^{\text{SM}}_{H} - \left(2\,\Delta\, c_{bH} - K_{bG}\,\Delta_{\text{non}}\, c_{bG}\right) \Gamma(H \to b\bar{b})^{\text{LO}}_{\text{SM}} \\ + 6\, K_{HG}\, c_{HG}\, \Gamma(H \to gg)^{\text{LO}}_{\text{SM}} \Big] , \end{aligned} \tag{5.20}$$

where $K_{bG} = 1.622$ represents the QCD correction to the partial decay width of $H \to b\bar{b}$ up to N$^3$LO related to $Q_{bG}$ . The specific value used follows from the semi-analytic formula (5.27). The term in $\Gamma^{\text{SMEFT}}_{H}$ proportional to

$$c_{HG} = \frac{v^2}{\Lambda^2} C_{HG} \tag{5.21}$$

represents the SMEFT correction to the partial decay width $H \to gg$ associated to the insertion of $Q_{HG}$. We show sample diagrams of this contribution in figure 5.2. This correction starts at NNLO since

$$\Gamma(H \to gg)^{\text{LO}}_{\text{SM}} = \frac{\alpha_s^2 m_H^3}{72\pi^3 v^2} , \tag{5.22}$$

which is obtained in the approximation of infinite top-quark mass. Higher-order effects proportional to $C_{HG}$ are encoded in an approximate way in the factor $K_{HG} = 1.844$ in equation 5.20 that include QCD corrections up to N$^4$LO [135]. Notice that equation (5.20) contains only corrections associated to the operators considered in (5.4).

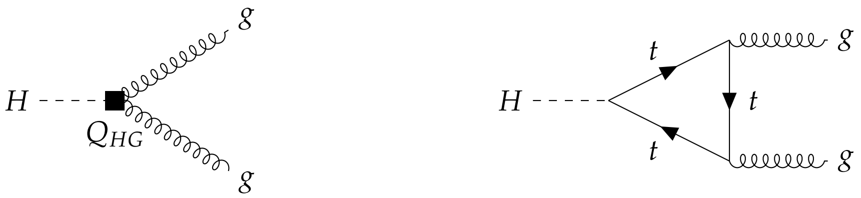


Figure 5.2.: Sample of interfering diagrams associated to SMEFT corrections proportional to $c_{HG}$. They contribute to the partial decay width $H \to gg$ needed in (5.20). The left diagram is the LO SMEFT result with the insertion of the $Q_{HG}$ operator, while the right diagram is the one-loop SM contribution mediated by a top-quark loop. In the latter case, there are also contributions mediated by light quarks but they have a negligible numerical impact.

## 5.4. Phenomenological results

### 5.4.1. Input and parameters

We now present phenomenological results for $pp \to ZH \to \ell^+\ell^- b\bar{b}$ for 13 TeV LHC collisions. We reach NNLO+PS accuracy in the SMEFT, considering the subset of dimension-six operators defined in (5.4). Our SM input parameters are taken from the PDG [140]:

$$\begin{aligned} G_F &= 1.166379 \times 10^{-5}\,\text{GeV}^{-2}\,, & \Gamma_Z &= 2.4952\,\text{GeV}\,, \\ m_W &= 80.379\,\text{GeV}\,, & m_H &= 125.09\,\text{GeV}\,, \\ m_Z &= 91.1876\,\text{GeV}\,, & \Gamma_H &= 4.1\,\text{MeV}\,. \end{aligned}$$

The EW coupling and the weak mixing angle are defined as:

$$\alpha = \sqrt{2} G_F m_W^2 \sin^2\theta_W/\pi\,, \qquad \sin^2\theta_W = 1 - m_W^2/m_Z^2\,.$$

The Higgs VEV is obtained using $v = 2^{-1/4}\, G_F^{-1/2}$. As for the $b$-quark mass, we use the $\overline{\text{MS}}$ mass $\overline{m}_b(\overline{m}_b) = 4.18\,\text{GeV}$ as input which leads to $\overline{m}_b(m_H) = 2.79\,\text{GeV}$ and $y_b(m_H) = \sqrt{2}\,\overline{m}_b(m_H)/v = 1.60 \cdot 10^{-2}$. We adopt the `NNPDF31_nnlo_as_0118` parton distribution functions [101] with $\alpha_s(m_Z^2) = 0.1180$, which corresponds to $\alpha_s(m_H^2) = 0.1127$. We shower our events using PYTHIA8 [87] with the Monash tune [102]. Hadronization, underlying event effects and QED showers are not included.

In the following, we will show results considering two possible scenarios:

$$1)\quad c_{\text{kin}} = 0,\quad c_{HG} = 0,\quad c_{3G} = 0,\quad c_{bH} \neq 0 \quad \text{and} \quad c_{bG} = 0\,, \tag{5.23}$$

$$2)\quad c_{\text{kin}} = 0,\quad c_{HG} = 0,\quad c_{3G} = 0,\quad c_{bH} = 0 \quad \text{and} \quad c_{bG} \neq 0\,. \tag{5.24}$$

We thus consider only the separate effect of $c_{bH}$ and $c_{bG}$. The current experimental constraints on these two operators at 95% CL are the following:

$$c_{bH} \in [-0.13, 0.20]\,, \tag{5.25}$$

$$c_{bG} \in [-438, 438]\,. \tag{5.26}$$

The bound on $c_{bH}$ is derived from a recent SMEFT fit [139], while the constraint on $c_{bG}$ comes from the analysis [138] of the transverse momentum spectrum of $Z$ production in association with $b$-jets as measured by ATLAS [141].

### 5.4.2. Inclusive $H \to b\bar{b}$ decay

In this section we present the inclusive decay width for $H \to b\bar{b}$ at N$^3$LO in the SMEFT in the limit of massless bottom quarks. We obtain

$$\Gamma(H \to b\bar{b})^{\mathrm{N^3LO}}_{\mathrm{SMEFT}} = \Bigg\{ (1 + 2\,c_{\mathrm{fac}}) \left[ 1 + \frac{\alpha_s}{\pi}\, 5.67 + \left(\frac{\alpha_s}{\pi}\right)^2 29.15 + \left(\frac{\alpha_s}{\pi}\right)^3 41.76 \right] \\ + \left(\frac{\alpha_s}{\pi}\right)^2 \frac{m_H^2}{3\,v^2} \left[ 1 + \frac{\alpha_s}{\pi}\, 17.32 \right] c_{bG} \Bigg\}\, \Gamma(H \to b\bar{b})^{\mathrm{LO}}_{\mathrm{SM}}\,, \tag{5.27}$$

where we set the renormalization scale to $\mu_R = m_H$. Notice that the $\mathcal{O}(\alpha_s^3)$ correction proportional to $c_{bG}$ is calculated by the author and collaborators in [3] for the first time. This contribution enhances significantly the non-factorizable corrections ($\sim$ 60%), which explains why encoding this correction in our NNLO+PS generator is crucial.

The $H \to b\bar{b}$ decay width shown in equation (5.27) allows us to study the impact of the operators defined in (5.4). We can consider the two scenarios in (5.23) and (5.24) and vary the coefficients $c_{bH}$ and $c_{bG}$ within their current experimental bounds (equations (5.25) and (5.26)). Looking at the shift induced by these operators, we can extrapolate a hierarchy in the SMEFT effects. More precisely:

$$\frac{\Gamma(H \to b\bar{b})^{\mathrm{N^3LO}}_{\mathrm{SMEFT}}}{\Gamma(H \to b\bar{b})^{\mathrm{N^3LO}}_{\mathrm{SM}}} - 1 \in \begin{cases} [-39,\, 26]\% & \text{in scenario 1 (5.23)}\,,\ \text{using (5.25)}\,, \\ [-6.3,\, 6.3]\% & \text{in scenario 2 (5.24)}\,,\ \text{using (5.26)}\,. \end{cases} \tag{5.28}$$

We observe that the impact of the $Q_{bG}$ operator is smaller by a factor $\mathcal{O}(5)$ than the factorizable corrections associated to the insertion of $Q_{bH}$.

### 5.4.3. Differential distributions

In this section we present our differential distributions employing the fiducial cuts reported in table 5.1. We stress that the cut on the transverse momentum of the $Z$ boson $p_{\mathrm{T},\ell^+\ell^-}$ is used in order to substantially reduce some backgrounds ($t\bar{t}$, single-top and diboson production [76]). Another source of background is given by $Z +$ jets production. We will comment on how our results vary if we modify the cut on the transverse momentum of the $Z$ boson considering the categories $p_{\mathrm{T},\ell^+\ell^-} \in [75, 150]$ GeV and $p_{\mathrm{T},\ell^+\ell^-} > 250$ GeV.

In the following plots we show our SMEFT results (red curve) considering either $c_{bH} = 0.15$ or $c_{bG} = 400$. The chosen numerical values are within the current experimental bounds (equations (5.25) and (5.26)). All the other Wilson coefficients are set to zero. We compare this computation with SM results as obtained in chapter 4. The error bars are given by the customary 7-point scale variation, correlating the scales in production and decay.

| fiducial cuts |
|---|
| $p_{\mathrm{T},\ell} > 15\,\mathrm{GeV}$, $\lvert\eta_\ell\rvert < 2.5$ |
| $75\,\mathrm{GeV} < m_{\ell^+\ell^-} < 105\,\mathrm{GeV}$, $p_{\mathrm{T},\ell^+\ell^-} \in [150, 250]\,\mathrm{GeV}$ |
| $\geq 2$ $b$-jets (anti-$k_T$ [104], R=0.4) |
| $p_{\mathrm{T},j} > 25\,\mathrm{GeV}$, $\lvert\eta_j\rvert < 2.5$ |

Table 5.1.: Fiducial phase space used in section 5.4.3.

We start our discussion from the invariant mass of the $b\bar{b}$ system. The two $b$-jets are reconstructed as the pair of $b$-jets whose invariant mass is closest to the mass of the Higgs boson. Including only factorizable corrections (left panel) clearly induces an overall rescaling of the SM distribution, which reaches 30% effect for the specific value of the Wilson coefficient we have chosen. A more interesting result is given by the inclusion of non-factorizable corrections (right panel) as it determines a distortion of the shape of the $m_{b\bar{b}}$ distribution. In fact, the SMEFT calculation receives important corrections below the Higgs threshold (up to 40% for $m_{b\bar{b}} \simeq 50\,\mathrm{GeV}$). This enhancement is related to the structure of the leading $Q_{bG}$ correction, which is associated to the contribution to $H \to b\bar{b}g$ depicted in the top left panel in figure 5.1. The interference of this contribution with the corresponding SM diagram produces an amplitude in which the gluon emission probability is isotropic in the phase space (the explicit result is reported in Appendix A of the original publication [3]). By contrast, in the SM the real emission of a gluon is strongly enhanced in the soft or collinear regions. As a result, configurations in which the energy-momentum of the Higgs boson is equally shared among the two bottom quarks and the gluon are more frequent in the SMEFT calculation, while in the SM we expect a $m_{b\bar{b}}$ distribution strongly peaked at the Higgs boson mass, as it is more likely that the two $b$-jets carry most of the energy-momentum of the Higgs. Moreover, we note that if we change the sign of the Wilson coefficients, the sign of the relative impact of the SMEFT corrections changes accordingly.

The shape of the SMEFT effects in the $m_{b\bar{b}}$ distribution depends on the jet radius $R$ used in the clustering algorithm for the identification of the $b$-jets. This feature can be observed comparing the right panel of figure 5.3 with both panels in figure 5.4. In these plots we show the $m_{b\bar{b}}$ distribution varying the radius in the clustering procedure. We show results for $R = 0.4$, $R = 0.7$ and $R = 1$. In general, the $m_{b\bar{b}}$ distribution is peaked at lower values when the radius increases. This same behaviour has been observed for different $p_{\mathrm{T},\ell^+\ell^-}$ categories. In fact, we analysed both $p_{\mathrm{T},\ell^+\ell^-} \in [75, 150]\,\mathrm{GeV}$ and $p_{\mathrm{T},\ell^+\ell^-} > 250\,\mathrm{GeV}$ and the same conclusions hold.

Another distribution that is particularly affected by SMEFT effects in presence of $c_{bG}$ is the invariant mass $m_{b\bar{b}j}$ of two $b$-jets and an extra jet. We reconstruct the $b\bar{b}j$ system taking two $b$-jets and an extra jet whose invariant mass is closest to the Higgs boson mass. In figure 5.5 we show the $m_{b\bar{b}j}$ distribution considering a radius of $R = 0.4$ (left panel) and $R = 0.7$ (right panel). SMEFT effects are strongly pronounced in the low $b\bar{b}j$-region because events

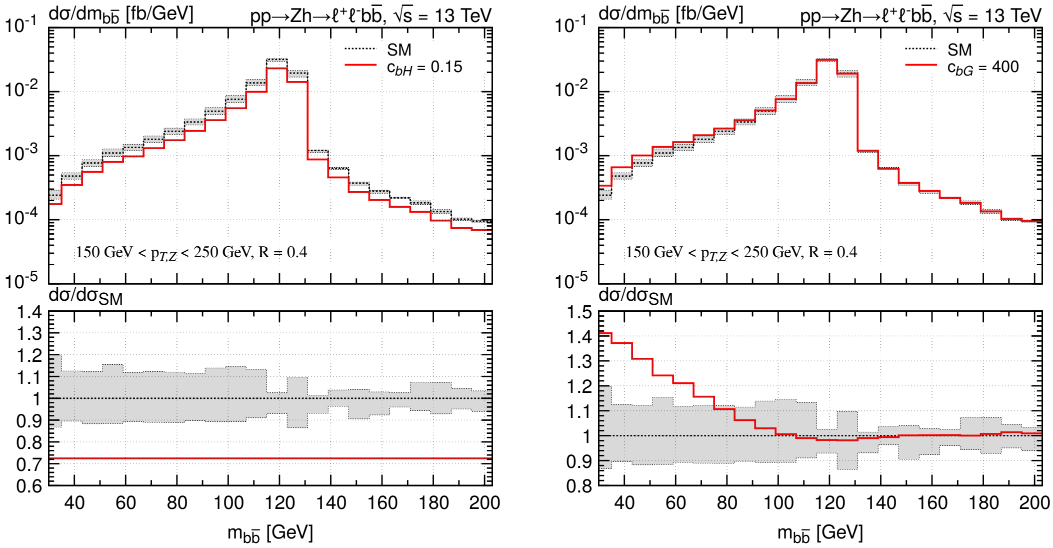


Figure 5.3.: Invariant mass of the two reconstructed $b$-jets in the fiducial setup reported in table 5.1 for $c_{bH} = 0.15$ (left panel) and $c_{bG} = 400$ (right panel). The SMEFT prediction is represented by the red curve. The SM result with its scale uncertainty band is shown in black and gray. The radius used in the clustering algorithm (anti-$k_T$) is $R = 0.4$.

with an extra jet are more likely in the SMEFT due to the insertion of the $Q_{bG}$ operator. Furthermore, we already observed that the largest corrections in the $m_{b\bar{b}}$ distributions arise at $\simeq$ 40 GeV (see right panel in figure 5.3 and both panels in 5.4). Since we apply a fiducial cut $p_{\mathrm{T},j} > 25\,\mathrm{GeV}$, we expect to observe an excess of events at $\simeq$ 60 GeV in the $m_{b\bar{b}j}$ distribution. Indeed, this effect is present in the left panel of 5.5. When comparing the two panels in 5.5 in which we increase the radius used in the clustering algorithm, we see that the SMEFT effects move toward larger values of $m_{b\bar{b}j}$: this is understood considering that a larger $R$ value means clustering more radiation in the same jet, which concretely produces a larger three-jet invariant mass on average. Notice that in the recent ATLAS analysis [142], the $m_{b\bar{b}j}$ distribution is already used for the construction of a multivariate discriminant for constraining $C_{bG}$. We believe that the shape and R-dependent effects presented in this section can increase our sensitivity to this operator, which is still weakly constrained.

We studied many observables other than $m_{b\bar{b}}$ and $m_{b\bar{b}j}$ for analyzing how the $Q_{bG}$ operator affects the process under study. We identified different distributions in which the impact of this operator is non-trivial, like the transverse momentum of the $Z$ boson ($p_{\mathrm{T},\ell^+\ell^-}$) and the transverse momentum of the two bottom quarks ($p_{\mathrm{T},bb}$). However, these effects cannot exceed the percent level, and thus these observables taken by themselves have limited constraining power. We thus suggest incorporating these observables in a multivariate discriminant that can enhance our overall sensitivity to $c_{bG}$.

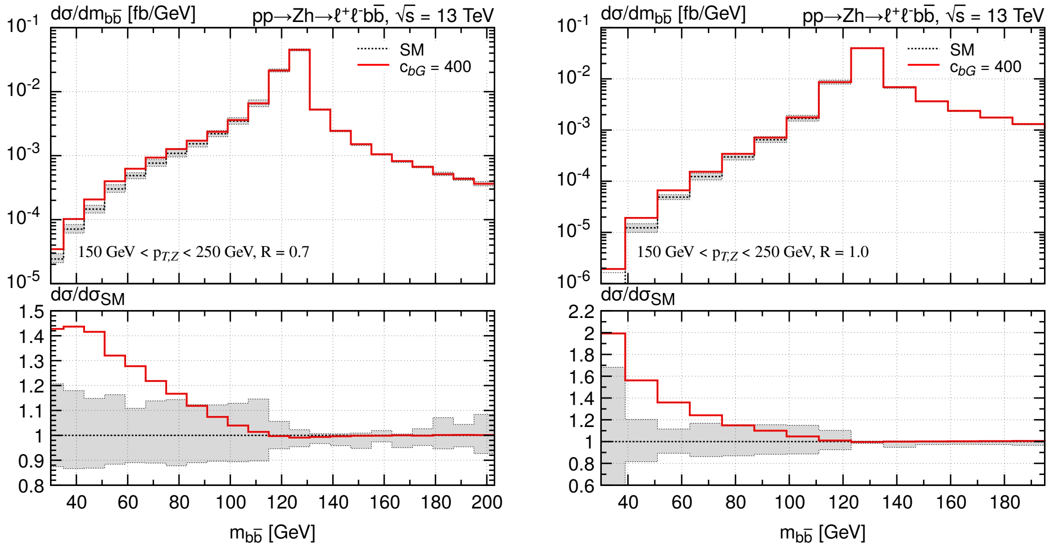


Figure 5.4.: Invariant mass of the two reconstructed $b$-jets in the fiducial setup reported in table 5.1 for $c_{bG} = 400$. The SMEFT prediction is represented by the red curve. The SM result with its scale uncertainty band is shown in black and gray. The radii used in the clustering algorithm (anti-$k_T$) are $R = 0.7$ (left panel) and $R = 1$ (right panel).

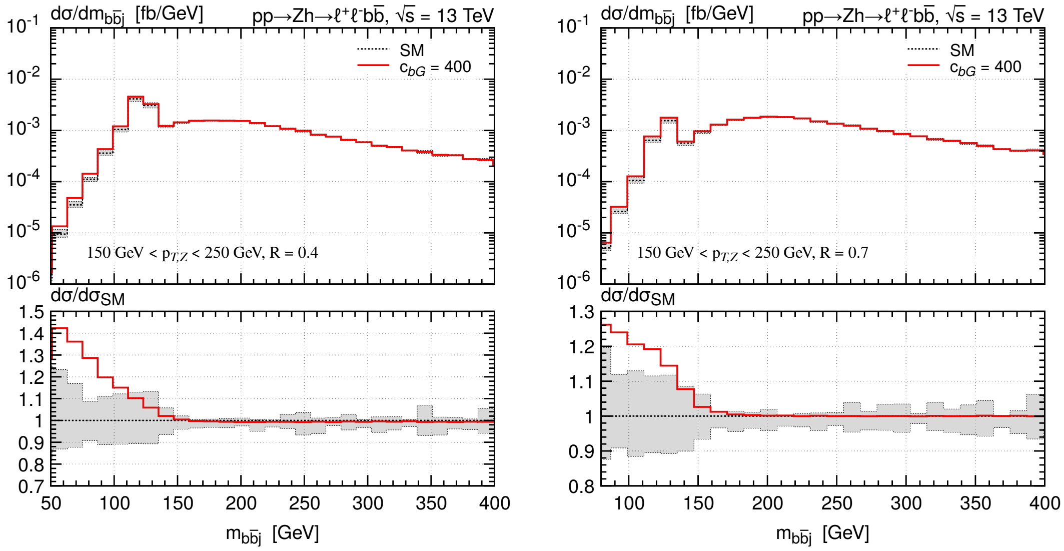


Figure 5.5.: Invariant mass of the $b\bar{b}j$ system in the fiducial setup reported in table 5.1 for $c_{bG} = 400$. The SMEFT prediction is represented by the red curve. The SM result with its scale uncertainty band is shown in black and gray. The radii used in the clustering algorithm (anti-$k_T$) are $R = 0.4$ (left panel) and $R = 0.7$ (right panel).

# 6. *WZ* production at NNLO QCD and NLO EW accuracy matched to parton showers

## 6.1. Motivation

In view of the increasing level of precision of current and future LHC experimental measurements, proposing theoretical calculations at the highest accuracy possible in both QCD and EW perturbation theory is crucial. In this context, the production of a pair of massive vector bosons $pp \to VV$ plays an important role, as experimental measurements for this kind of process are precision tests of the SM. In fact, $VV$ production provides direct access to the trilinear gauge couplings, which may be modified by new physics effects, and it thus probes the gauge symmetry structure of the EW sector and its interplay with the scalar one. The aim of this chapter is to present a NNLO QCD and NLO EW accurate calculation matched to parton showers for $pp \to W^\pm Z \to \ell'^\pm \nu_{\ell'} \ell^+ \ell^-$. We consider the specific case of $W^\pm Z$ production because it has a large cross section and a clean experimental signature if we consider the fully leptonic decay (the signature is given by three charged leptons and missing energy). However, we highlight that the strategy we present in this chapter is fully general and it can be applied to any process of interest, provided the correct matrix elements are available.

A combination of QCD and EW corrections is non-trivial, especially when matched to parton showers. Thus, we first present how to combine NNLO QCD and NLO EW contributions at fixed order reviewing the discussion presented in [143]. To this end, we define the following notation

$$\mathrm{d}\sigma_{\text{NNLO QCD}} = \mathrm{d}\sigma_{\text{LO}}\left(1+\delta_{\text{QCD}}\right), \qquad \mathrm{d}\sigma_{\text{NLO EW}} = \mathrm{d}\sigma_{\text{LO}}\left(1+\delta_{\text{EW}}\right), \tag{6.1}$$

where the higher-order corrections with respect to the LO contribution are encoded into $\delta_{\text{QCD}}$ and $\delta_{\text{EW}}$ for QCD and EW calculations, respectively. The combination of QCD and EW corrections can be performed through two main schemes:

- **Additive scheme**

  Higher-order QCD and EW corrections can be combined in a purely additive way via

  $$\mathrm{d}\sigma_{\text{NNLO QCD+EW}} = \mathrm{d}\sigma_{\text{NNLO QCD}} + \mathrm{d}\sigma_{\text{NLO EW}} - \mathrm{d}\sigma_{\text{LO}} = \mathrm{d}\sigma_{\text{LO}}\left(1+\delta_{\text{QCD}}+\delta_{\text{EW}}\right), \tag{6.2}$$

  in which we simply sum the NNLO QCD and the NLO EW computations, removing the LO contribution to avoid double counting. Considering the specific case of $W^\pm Z$ production, using this combination procedure we generate terms of order $\mathcal{O}(\alpha^4)$ (LO), $\mathcal{O}(\alpha^4\alpha_s)$ (NLO QCD), $\mathcal{O}(\alpha^4\alpha_s^2)$ (NNLO QCD) and $\mathcal{O}(\alpha^5)$ (NLO EW).

- **Multiplicative scheme**

  Higher-order QCD and EW corrections can be combined in a factorized scheme

  $$\mathrm{d}\sigma_{\text{NNLO QCD}\times\text{EW}} = \mathrm{d}\sigma_{\text{NNLO QCD}} \times \frac{\mathrm{d}\sigma_{\text{NLO EW}}}{\mathrm{d}\sigma_{\text{LO}}} = \mathrm{d}\sigma_{\text{LO}}\left(1+\delta_{\text{QCD}}\right)\left(1+\delta_{\text{EW}}\right), \tag{6.3}$$

  where we supplement the NNLO QCD computation with an EW K-factor given by the ratio of the NLO and the LO computations. Using this combination scheme we generate mixed QCD-EW contributions. Considering the specific case of $W^{\pm}Z$ production, we obtain extra terms of order $\mathcal{O}(\alpha^5\alpha_s)$ and $\mathcal{O}(\alpha^5\alpha_s^2)$ when compared to the additive combination.

The multiplicative scheme should be seen as superior to the additive combination (6.2) because of the factorization of QCD and EW effects at high energies. More precisely, EW effects are dominated by EW Sudakov logarithms at high energies, while the dominant QCD effects are related to scales much lower than the hard scale so that QCD and EW effects factorize. However, $VV$ production presents topologies for which this assumption is violated and neither a multiplicative nor an additive scheme is suitable for performing the QCD-EW combination. These topologies are associated with so-called *giant K-factors*.

Giant K-factors arise from configurations in which the gap between the hardness of the two vector bosons is large: $p_{\text{T},V_1} \gg p_{\text{T},V_2}$. They are thus pure NLO effects because they are associated with regions of the phase space that are forbidden at LO. In fact, when a hard vector boson $V_1$ is present at LO, the second vector boson $V_2$ is hard as well due to energy-momentum conservation, $p_{\text{T},V_1} = p_{\text{T},V_2}$. However, at NLO most of the recoiling energy can be absorbed by the extra jet $p_{\text{T},V_1} \simeq p_{\text{T},J}$, while the second vector boson remains soft $p_{\text{T},V_2} = \mathcal{O}(m_W)$. Anytime a process is characterized by a topology with a hard vector boson, a hard jet and a soft second vector boson, we can integrate over the soft vector boson phase space [144]:

$$\mathrm{d}\sigma^{V(V)J} \propto \mathrm{d}\sigma^{VJ}_{\text{LO}} \frac{\alpha}{2\pi} \log^2\left(\frac{Q^2}{m_W^2}\right). \tag{6.4}$$

The previous equation shows that we can factorize the cross section for a hard $VJ$ process with a soft second vector boson into a hard $VJ$ cross section at LO ($\mathrm{d}\sigma^{VJ}_{LO}$) multiplied by a radiative correction that represents the emission probability of the soft vector boson. Given that $\sigma^{VJ}_{\text{LO}}/\sigma^{VV}_{\text{LO}} \propto \alpha_s/\alpha$, the radiative correction to the hard $VV$ process is

$$\frac{\mathrm{d}\sigma^{V(V)J}}{d\sigma^{VV}_{\text{LO}}} \propto \alpha_s \log^2\left(\frac{Q^2}{m_W^2}\right), \tag{6.5}$$

which becomes large already at the TeV scale (e.g. it is a factor of 3 when $Q = 1\,\text{TeV}$).

We stress one more time that giant K-factors are pure NLO effects which are different from enhanced QCD logarithms associated with soft and collinear emissions. While the former effects are related to the opening of a new channel at NLO accuracy (the hard $pp \to V(V)J$ channel), the latter appear to all orders in perturbation theory. Starting from NNLO accuracy, calculations are free from giant K-factors.

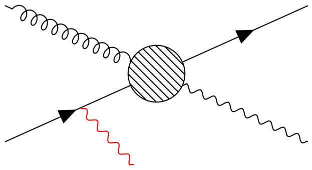


Figure 6.1.: Sample Feynman diagram associated with giant K-factors at NLO QCD. The blob represents the hard scattering at scale $Q \gg m_W$ while the red vector boson is the subleading one which generates EW soft and collinear logarithms.

The appearance of giant K-factors can spoil the convergence of the perturbative series and it creates issues in combining QCD and EW higher-order contributions. In fact, a multiplicative scheme is not suitable in this case: NLO EW effects are dominated by Sudakov logarithms to the hard $VV$ process, while giant QCD K-factors are driven by soft EW boson radiation in the $VJ$ process. There is no factorization of QCD and EW effects so that a multiplicative scheme is not motivated. When giant K-factors are present, the average between the multiplicative and the additive schemes can give us a pragmatic prediction of our result, and the difference between the two provides an estimate of error bars associated with missing higher-order effects. Notice that configurations associated with giant K-factors can be avoided by means of appropriate (dynamic) jet-veto that can suppress configurations associated with hard $VJ$ topologies.

$W^{\pm}Z$ production was extensively studied in the past years. Calculations at NLO accuracy in QCD were presented in [145, 146, 147, 148, 149] and results for polarized $W^{\pm}Z$ production are available in the double-pole approximation [150]. NLO QCD results for $W^{\pm}Z$+jet were obtained in [151]. At NNLO accuracy, inclusive cross sections were presented in [152] and an extension to fully differential predictions can be found in [153]. NNLO computations are nowadays publicly available in Matrix [154] and MCFM [155]. As far as EW corrections are concerned, NLO results were presented for on-shell production in [156, 157] and including off-shell decays in [158]. A combination of NNLO QCD and NLO EW computations was performed through Matrix+OpenLoops as presented in [143]. Results matched to parton showers were presented at NLO accuracy in [159, 160].

In this chapter, we extend the problem of the combination of QCD and EW higher-order contributions to the matching with parton showers, considering the specific case of $W^{\pm}Z$ production. The main issue in reaching this accuracy is defining a consistent way of combining QCD and EW calculations with QCD and/or QED parton showers without spoiling the formal accuracy of the fixed-order computation and without including any double counting.

This chapter is organised as follows: in section 6.2 we describe our calculation and the possible matching schemes between NNLO QCD and NLO EW computations matched with parton showers; in section 6.4 we present a phenomenological analysis for 13 TeV collisions at the LHC commenting in detail on the differences among the different matching schemes and presenting our best prediction. In the same section, we show a comparison with recent ATLAS data [161] as well. The results presented in this chapter were originally presented by the author and collaborators in [4].

## 6.2. Outline of the calculation

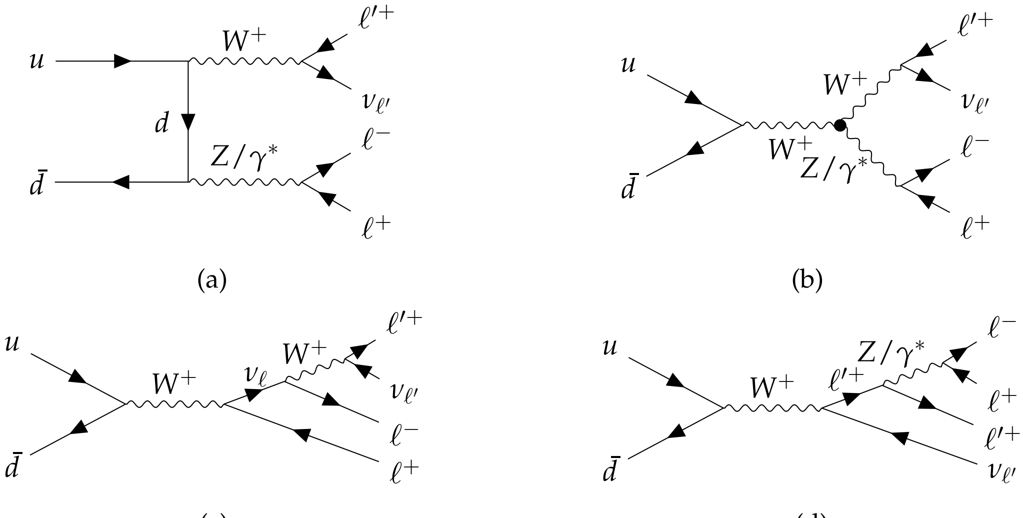


Figure 6.2.: Feynman diagrams for $W^+Z$ production at LO. Panel (a) is the $t$-channel contribution, panel (b) is the $s$-channel contribution and panels (c,d) are DY-type channels.

We consider the process

$$pp \to \ell'^{\pm}\nu_{\ell'}\ell^+\ell^- + X\,, \tag{6.6}$$

where $\ell$ and $\ell'$ represent any combination of massless leptons ($\ell, \ell' \in \{e, \mu, \tau\}$). We study both different flavour ($\ell \neq \ell'$) and same flavour ($\ell = \ell'$) channels. Sample Feynman diagrams for this process are presented in figure 6.2, in which we show LO contributions ($\mathcal{O}(\alpha^4)$).

As far as NNLO QCD corrections are concerned, we include contributions up to $\mathcal{O}(\alpha^4\alpha_s^2)$ and we note that no gluon-gluon contributions are possible due to charge conservation. NNLO corrections are expected to be large because of the *radiation zero effect* at LO [162] for which leading helicity amplitudes vanish in some kinematic regions. As a consequence, the LO calculation is not reliable and higher-order contributions are particularly large (e.g. NNLO corrections are of the order of 10–15% [152]).

The NLO EW contribution is of order $\mathcal{O}(\alpha^5)$. Virtual corrections are $q\bar{q}$ one-loop contributions involving $W$, $Z$ and Higgs bosons, photons and fermions in the loop, including heavy quarks. Real corrections are of pure QED type (photon emissions). Photon-induced contributions $\gamma\gamma$ are not present at this order, while photon-quark $\gamma q$ contributions are in principle allowed. The latter contributions have been neglected because of the suppressed photon flux in the proton, which is accompanied by an extra power of the electroweak coupling $O(\alpha)$ times a collinear logarithm $L$.[1]

---

[1] We note that $\gamma q$ contributions are potentially large when the initial-state photon $\gamma$ couples directly to a $W$-

We reach NNLO QCD and NLO EW accuracy matched to parton showers using an a-posteriori combination. Our strategy proceeds in three main steps: first, we separately generate events at NNLO QCD and NLO EW accuracy within the POWHEG-BOX-RES framework [88]. Secondly, we shower the events through PYTHIA8 [87] applying a dedicated veto procedure. Lastly, we combine QCD and EW results at the level of differential distributions using appropriate combination schemes.

### 6.2.1. Generation of events

We separately generate NNLO QCD and NLO EW accurate results within the POWHEG-BOX-RES framework.

As for the QCD computation, we adopt the MiNNLO$_{\text{PS}}$ procedure. Since no $W^{\pm}Z$ generator was available in the POWHEG-BOX-RES, we first implemented the corresponding $W^{\pm}ZJ$ code and then we upgraded it to NNLO accuracy in $W^{\pm}Z$ production through the MiNNLO$_{\text{PS}}$ method. The matrix elements for tree-level and one-loop corrections are obtained from OpenLoops [163, 164, 165] while the two-loop amplitude is provided by VVamp [166] through the interface to Matrix [154]. Given the CPU-demanding computation, the two-loop correction is included only at the level of the generation of events (stage 4) in POWHEG, setting appropriately the `run_mode` option in the input card as described in [70]. For the calculation of the $D$ terms in the MiNNLO$_{\text{PS}}$ formula (3.45), we use HOPPET [74] for a fast evaluation of the convolutions with PDFs. Furthermore, we compute polylogarithms in the collinear coefficient functions using hplog [93]. To avoid spurious higher-order logarithmic terms in the MiNNLO$_{\text{PS}}$ formula (3.45), we adopt a modified logarithm (3.53) that smoothly turns off NNLO corrections in the region $p_{\text{T}} > m_{W^{\pm}Z}$, where $p_{\text{T}}$ and $W^{\pm}Z$ are the transverse momentum and the invariant mass of the $W^{\pm}Z$ colour singlet [71]. At high $p_{\text{T}}$, we set the scales of the calculation through the `largeptscales 1` option so that equation (3.58) is used. At small $p_{\text{T}}$, the default MiNNLO$_{\text{PS}}$ prescription in the choice of the scales is employed. Furthermore, we set $Q_0 = 0$ GeV in (3.57) and we freeze the strong coupling and the PDFs at 0.8 GeV. We also switch on the POWHEG option doublefsr 1 [94], so that $q \to qg$ and $g \to q\bar{q}$ splittings are treated symmetrically for the definition of the starting scale of the shower.

For the computation of the NLO EW calculation, we developed a code analogous to the one constructed in [159]. More precisely, we implemented a $W^{\pm}Z$ generator within the POWHEG-BOX-RES framework. This code can produce NLO QCD, NLO EW and NLO QCD + NLO EW results matched to QCD and QED parton showers for all $pp \to VV$ processes. All the needed matrix elements are taken from OpenLoops [163, 164, 165]. Also in this case, we set `doublefsr 1`.

### 6.2.2. Matching with parton showers

The parton shower is generated by PYTHIA8 [87]. We keep all the standard settings to their default value, in particular for the recoil scheme. The matching of NNLO QCD and NLO EW

---

boson propagator in the $t$-channel [158]. These contributions can be implemented separately and added to higher-order $q\bar{q}$ QCD corrections in a purely additive way.

accurate results with QCD and QED parton showers requires a dedicated veto procedure. Our strategy[2] consists in letting both the QCD and QED showers radiate in the entire kinematically allowed phase space and then checking that the shower history is consistent with the emission generated by POWHEG at LHE level. More precisely, we activate the following options in PYTHIA8

```
pythia.readString("SpaceShower:pTmaxMatch = 2"),
pythia.readString("TimeShower:pTmaxMatch = 2"),
```

which set the starting scales of both space- and time-like showers to the partonic energy of the event. Once the event is showered, we apply an a-posteriori veto as follows:

- When generating NNLO QCD events, we restrict the QCD radiation produced by the shower, as default in POWHEG. QED radiation from the shower remains unconstrained so that the entire phase space allowed by the kinematics is covered. The veto procedure thus consists in scanning all the QCD emissions produced by PYTHIA8 and storing the hardest transverse momentum $p_{\mathrm{T}}^{\max}$. We now compare $p_{\mathrm{T}}^{\max}$ with the hardness (`scalup`) of the QCD emission generated by POWHEG at LHE level. The value of `scalup` is read from the event file. We accept the event only if $p_{\mathrm{T}}^{\max} < \texttt{scalup}$. If this requirement is not met, we try to reshower the event again. After 1000 failed attempts, the event is rejected.
- As for the generation of NLO EW predictions, we apply the same logic: QED emissions must be restricted to avoid double counting, while QCD radiation should cover the entire allowed phase space. QED radiation can be produced by the shower both as initial-state radiation (ISR) and as final-state radiation (FSR). In the latter case, it can come from the decay of both the $W$ (FSR-W) and the $Z$ (FSR-Z) bosons. We thus define three different starting scales, according to the region in which the QED emission is produced. For doing so, we activate the multiple-radiation scheme option (allrad 1 in the input card) of POWHEG. In this way, we keep track of all the possible QED emissions generated by POWHEG in the three singular regions of the process. The transverse momenta of these (up to) three emissions define the starting scales $\texttt{scalup}_i$ in the different regions ($i \in$ {ISR, FSR-W, FSR-Z}). More precisely, in the case of FSR we calculate the transverse momentum of the photon with respect to the lepton emitter in the centre-of-mass frame of the mother resonance. If no photon is produced by POWHEG in a certain region, we set `scalup` to a default IR cutoff ($10^{-3}$ GeV). We then scan all the QED emissions generated by PYTHIA8, we store the hardest transverse momenta $p_{\mathrm{T},i}^{\max}$ for each region and then we check whether the requirement $p_{\mathrm{T},i}^{\max} < \texttt{scalup}_i$ is fulfilled for each region. If yes, we accept the event, otherwise we try to shower the event again. After 1000 unsuccessful attempts, the event is rejected.

### 6.2.3. Combination of QCD and EW results

Once the NNLO QCD and NLO EW results have been separately showered according to the procedure defined in 6.2.2, we can combine them at the level of differential distributions.

---

[2] This veto procedure is inspired by what is done in Appendix D in [78].

In order to make the notation easier, we will refer to these results as $\mathrm{NNLO_{QCD}}$+PS and $\mathrm{NLO_{EW}}$+PS, respectively. Moreover, we introduce

$$\mathrm{(N)NLO}_{\mathrm{X}}^{(\mathrm{Y})_{\mathrm{PS}}} \quad \text{with} \quad \begin{cases} X \in \{\mathrm{QCD,\ EW}\}\,, \\ Y \in \{\mathrm{QCD,\ QED,\ QCD\ and\ QED}\}\,, \end{cases} \tag{6.7}$$

which refers to the (N)NLO calculation in X perturbation theory matched to Y parton showers. We also introduce a generic term for the higher-order correction to the LO computation defined as

$$\delta\mathrm{N(N)LO_X}^{(\mathrm{Y})_{\mathrm{PS}}} = \mathrm{N(N)LO}_{\mathrm{X}}^{(\mathrm{Y})_{\mathrm{PS}}} - \mathrm{LO}_{\mathrm{X}}^{(\mathrm{Y})_{\mathrm{PS}}}\,, \tag{6.8}$$

and a multiplicative K-factor:

$$\text{K-}\mathrm{N(N)LO}_{\mathrm{X}}^{(\mathrm{Y})_{\mathrm{PS}}} = \mathrm{N(N)LO}_{\mathrm{X}}^{(\mathrm{Y})_{\mathrm{PS}}} / \mathrm{LO}_{\mathrm{X}}^{(\mathrm{Y})_{\mathrm{PS}}}\,. \tag{6.9}$$

At fixed order, QCD and EW calculations can be combined through the additive and multiplicative schemes presented in equations (6.2) and (6.3). When considering the matching with parton showers, this combination should not spoil the formal accuracy of the fixed-order computation and should not introduce any source of double counting. We thus identify the following possible combinations of $\mathrm{NNLO_{QCD}}$+PS and $\mathrm{NLO_{EW}}$+PS results:

Additive schemes:

$$1.\ \mathrm{NNLO}_{\mathrm{QCD}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}} + \delta\mathrm{NLO}_{\mathrm{EW}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}} = \mathrm{NNLO}_{\mathrm{QCD+EW}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}\,, \tag{6.10}$$

$$2.\ \mathrm{NNLO}_{\mathrm{QCD}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}} + \delta\mathrm{NLO}_{\mathrm{EW}}^{(\mathrm{QED})_{\mathrm{PS}}}\,, \tag{6.11}$$

$$3.\ \mathrm{NLO}_{\mathrm{EW}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}} + \delta\mathrm{NNLO}_{\mathrm{QCD}}^{(\mathrm{QCD})_{\mathrm{PS}}}\,. \tag{6.12}$$

Multiplicative schemes:

$$4.\ \mathrm{NNLO}_{\mathrm{QCD}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}} \times \text{K-}\mathrm{NLO}_{\mathrm{EW}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}} = \mathrm{NNLO}_{\mathrm{QCD\times EW}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}\,, \tag{6.13}$$

$$5.\ \mathrm{NNLO}_{\mathrm{QCD}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}} \times \text{K-}\mathrm{NLO}_{\mathrm{EW}}^{(\mathrm{QED})_{\mathrm{PS}}}\,, \tag{6.14}$$

$$6.\ \mathrm{NLO}_{\mathrm{EW}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}} \times \text{K-}\mathrm{NNLO}_{\mathrm{QCD}}^{(\mathrm{QCD})_{\mathrm{PS}}}\,, \tag{6.15}$$

$$7.\ \mathrm{NNLO}_{\mathrm{QCD}}^{(\mathrm{QCD})_{\mathrm{PS}}} \times \text{K-}\mathrm{NLO}_{\mathrm{EW}}^{(\mathrm{f.o.})}\,. \tag{6.16}$$

We introduced the short-hand notations $\mathrm{NNLO}_{\mathrm{QCD+EW}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}$ and $\mathrm{NNLO}_{\mathrm{QCDxEW}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}$ for what will be referred to as *default* additive and multiplicative schemes (first and fourth combinations). We also introduced the fixed-order NLO EW K-factor

$$\text{K-}\mathrm{NLO}_{\mathrm{EW}}^{(\mathrm{f.o.})} = \mathrm{NLO}_{\mathrm{EW}}^{(\mathrm{f.o.})} / \mathrm{LO}^{(\mathrm{f.o})}\,, \tag{6.17}$$

where we do not perform any matching with the parton shower (it is obtained via MATRIX+OPENLOOPS).

All the presented scheme have the same formal accuracy and they differ only for higher-order terms (note that the last combination does not include the QED shower). We will analyse these combination schemes in detail and we will discuss their ability to describe relevant kinematic regions of distributions.

## 6.3. Validation of the calculation

In this section we validate both the NNLO$_{\text{QCD}}$+PS and the NLO$_{\text{EW}}$+PS implementations through a comparison with fixed-order calculations. We present results for $pp \to \mu^+\nu_\mu e^+e^-$ at 13 TeV centre-of-mass energy, but the same conclusions hold for any decay channel. The validation is performed at LHE level. The fixed-order results for both NNLO QCD and NLO EW computations are obtained through MATRIX+OPENLOOPS.

### 6.3.1. Input and parameters

We use the `NNPDF31_nnlo_as_0118_luxqed` [167, 168, 169] PDF set, which corresponds to $\alpha_s(m_Z^2) = 0.118$. Our input parameters are taken from the PDG [9]:

$$\begin{aligned}
G_F &= 1.16639 \times 10^{-5}\,\text{GeV}^{-2}\,, & m_H &= 125\,\text{GeV}\,,\\
m_W &= 80.385\,\text{GeV}\,, & \Gamma_H &= 4.07\,\text{MeV}\,,\\
\Gamma_W &= 2.0854\,\text{GeV}\,, & m_t &= 173.2\,\text{GeV (on-shell)}\,,\\
m_Z &= 91.1876\,\text{GeV}\,, & \Gamma_t &= 1.347878\,\text{GeV}\,.\\
\Gamma_Z &= 2.4952\,\text{GeV}\,, & &
\end{aligned}$$

We adopt the complex-mass scheme [170, 171] and EW parameters are determined through the $G_\mu$ scheme [165]:

$$\alpha_{G_\mu} = \frac{\sqrt{2}}{\pi}\, G_F |(m_W^2 - i\Gamma_W m_W)\sin^2\theta_W|\,, \qquad \cos^2\theta_W = \frac{m_W^2 - i\Gamma_W m_W}{m_Z^2 - i\Gamma_Z m_Z}\,.$$

NNLO$_{\text{QCD}}$+PS results are obtained using the MINNLO$_{\text{PS}}$ method, so the central scales are set using the standard MINNLO$_{\text{PS}}$ procedure. As for the NLO$_{\text{EW}}$+PS calculation, we use the following central scales

$$\mu_{\text{R}} = \mu_{\text{F}} = \frac{1}{2}\left(\sqrt{m_{e^+e^-}^2 + p_{\text{T},e^+e^-}^2} + \sqrt{m_{\mu\nu_\mu}^2 + p_{\text{T},\mu\nu_\mu}^2}\right)\,, \tag{6.18}$$

where $m_{e^+e^-}$ and $p_{\text{T},e^+e^-}$ ($m_{\mu\nu_\mu}$ and $p_{\text{T},\mu\nu_\mu}$) are the invariant mass and the transverse momentum of the reconstructed $Z$ boson ($W$ boson). We estimate error bars using the customary 7-point scale variation, keeping $0.5 \le \mu_R/\mu_F \le 2.0$. In the fixed-order computations, we use the same central renormalization and factorization scales as defined in (6.18).

Leptons are obtained using a dressing procedure: for every bare lepton $\ell^b$ we construct a cone of radius $R = 0.1$ around it. All the photons inside this cone ($\Delta R_{\ell^b\gamma} = \sqrt{\Delta\phi_{\ell^b\gamma}^2 + \Delta\eta_{\ell^b\gamma}^2} < 0.1$) are clustered together with the bare lepton, defining the physical lepton $\ell$ with momentum $p_\ell = p_{\ell^b} + p_\gamma$. The clustering procedure starts from the smallest $\Delta R_{\ell^b\gamma}$ distance among all the possible lepton-photon couples. Once a photon is recombined with a lepton, it is removed from the final-state particles of the process.

We consider an inclusive setup (`inclusive setup`) in which we apply only a cut on the $Z$ boson mass:

$$66\,\text{GeV} < m_{e^+e^-} < 116\,\text{GeV}\,. \tag{6.19}$$

This mass window is needed to avoid the photon singularity.

### 6.3.2. NNLO QCD

Figure 6.3 shows the comparison between NNLO$_{\mathrm{QCD}}$+PS results at LHE level (blue curve) and fixed-order results at NNLO QCD (red curve). The former are obtained through the MiNNLO$_{\mathrm{PS}}$ procedure as explained in the previous sections while the latter are obtained using Matrix+OpenLoops. We consider the `inclusive setup`.

We show the following differential observables: the difference in the azimuthal angle between the two leading charged leptons ($\Delta\phi_{\ell\ell}$), the rapidity of the leading charged lepton ($y_{\ell_1}$), the invariant mass of the three charged leptons ($m_{3\ell}$), the invariant mass ($m_{\mu\nu_\mu}$) and the transverse momentum ($p_{\mathrm{T},\mu\nu_\mu}$) of the reconstructed $W$ boson, and the missing transverse momentum ($p_{\mathrm{T,miss}}$). For all of them, we see an excellent agreement between the MiNNLO$_{\mathrm{PS}}$ and the fixed-order results. In the two calculations, error bars have a similar size. In this comparison, we do not expect a one-to-one agreement, as the two computations differ for the treatment of higher-order corrections and in the scale settings.

This comparison validates the MiNNLO$_{\mathrm{PS}}$ implementation.

### 6.3.3. NLO EW

Figure 6.4 shows the comparison between NLO$_{\mathrm{EW}}$+PS predictions at LHE level (blue curve) and fixed-order results at NLO EW (red curve). The former are obtained through POWHEG as explained in the previous sections while the latter are obtained using Matrix+OpenLoops. We consider the `inclusive setup`.

In this comparison, we expect a one-to-one agreement in the two predictions apart from small numerical fluctuations in the tails of distributions. In fact, the two computations are identical except for the POWHEG Sudakov in the NLO$_{\mathrm{EW}}$+PS results that correctly generates a photon emission for the matching with a parton shower.

We show the same distributions already presented in the validation of the NNLO QCD generator. For all of them, we see an excellent agreement, both in the central value and in the size of error bars.

This comparison validates the NLO EW POWHEG implementation.

## 6.4. Phenomenological results

In this section we present phenomenological results for $W^{\pm}Z$ production at NNLO QCD and NLO EW matched to parton showers, as originally presented in [4]. For simplicity, we again restrict ourselves to the specific case:

$$pp \to \mu^+\nu_\mu e^+e^- \,. \tag{6.20}$$

### 6.4.1. Settings

Our input and settings are identical to the ones used in the validation, so we refer the reader to section 6.3.1.

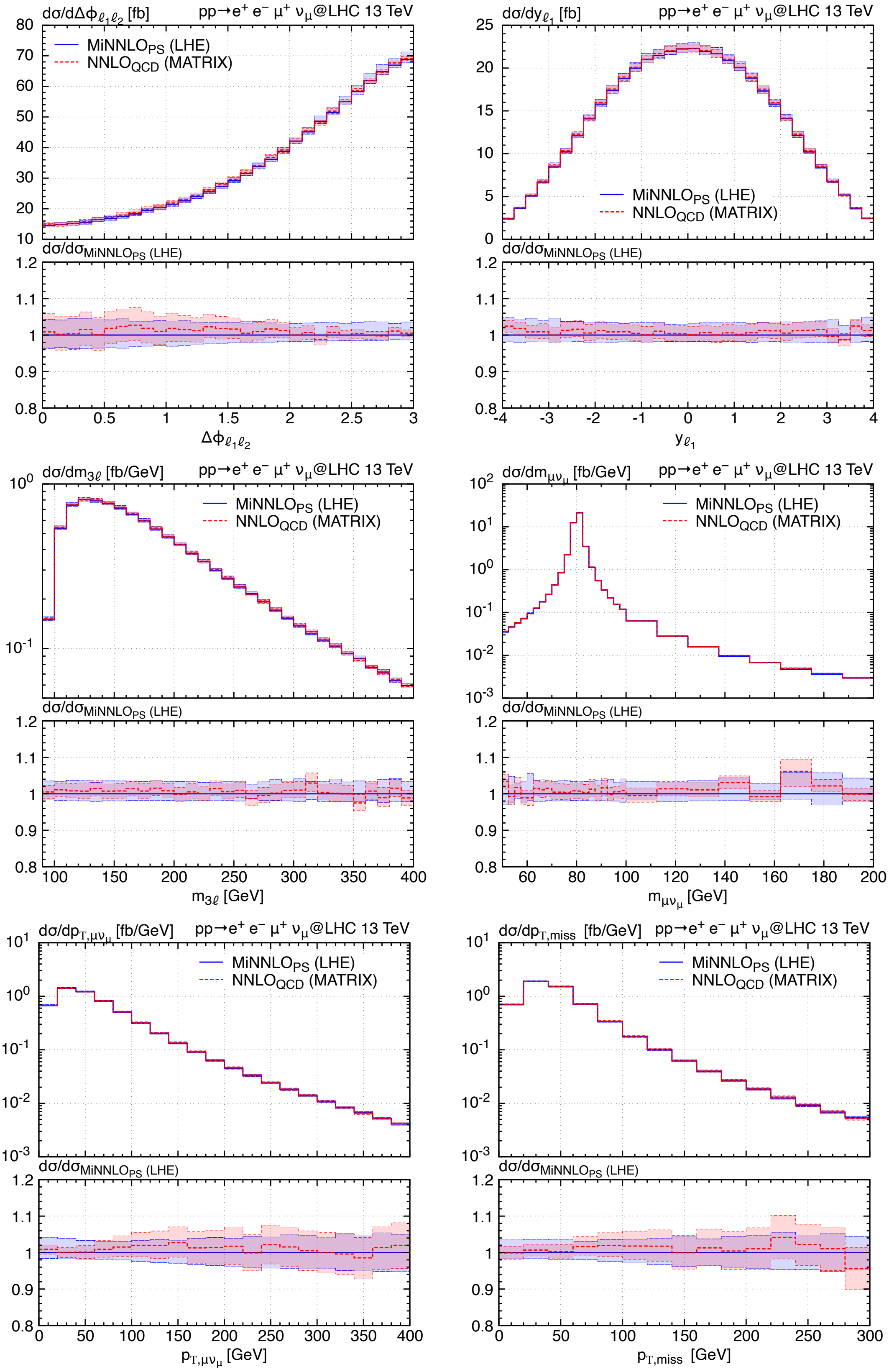


Figure 6.3.: $W^+Z$ predictions at NNLO QCD accuracy obtained using the MiNNLO$_{\rm PS}$ method (blue) and obtained at fixed-order using Matrix+OpenLoops (red) in the `inclusive setup`.

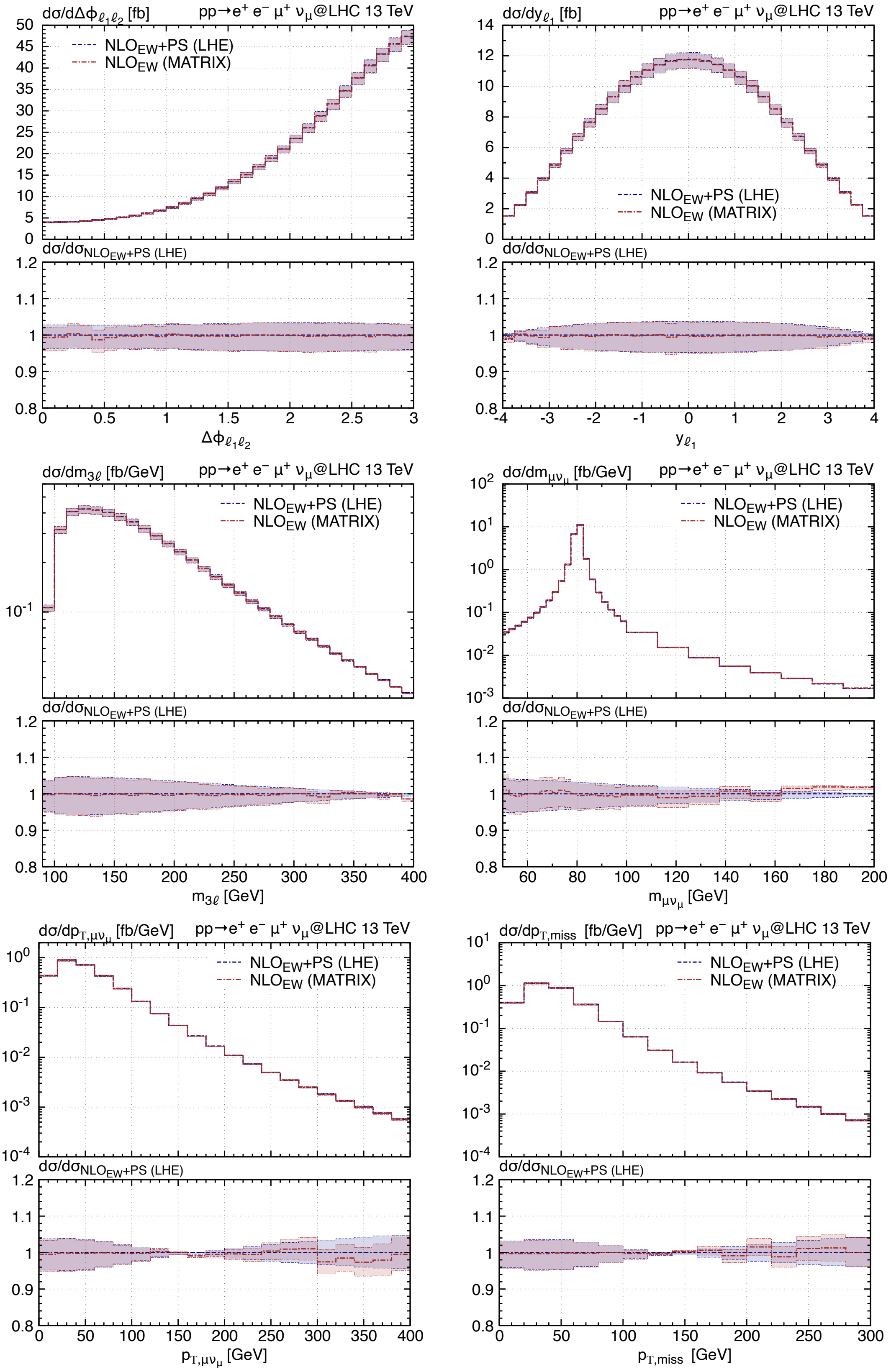


Figure 6.4.: $W^+Z$ predictions at NLO EW accuracy obtained using POWHEG (blue) and obtained at fixed-order using Matrix+OpenLoops (red) in the `inclusive setup`.

When combining QCD and EW results, we always correlate the scales in the two computations. We thus note that, when using a multiplicative scheme, the EW K-factor is almost scale independent, up to corrections of order $\mathcal{O}(\alpha)$ when varying the factorization scale.

The parton shower is generated by PYTHIA8 [87] with the Monash 2013 tune [102]. We set the option `TimeShower:QEDshowerByOther` to `off`, preventing resonances to emit photons, in order not to spoil the NLO resonance structure as generated by POWHEG. Moreover, also `TimeShower:QEDshowerByGamma` is set to `off`, preventing photons to split in leptons and quarks. Hadronization effects and multi-particle interactions (MPI) are considered only in section 6.4.3, where we perform a comparison with recent ATLAS data.

Results are presented using two different definitions of the fiducial phase space, as defined in table 6.1. The first setup is the inclusive one (`inclusive setup`) already presented in equation (6.19), while the second setup is taken from the ATLAS analyses of [172] and [161].

| `inclusive setup` |
|---|
| $66\,\text{GeV} < m_{e^+e^-} < 116\,\text{GeV}$ |

| `fiducial setup` |
|---|
| $\lvert m_{e^+e^-} - m_Z \rvert < 10\,\text{GeV}$ |
| $p_{T,e^\pm} > 15\,\text{GeV}, \quad p_{\text{T},\mu} > 20\,\text{GeV},$ |
| $\lvert \eta_\ell \rvert < 2.5, \quad m_{\text{T,W}} > 30\,\text{GeV},$ |
| $\Delta R_{e^+e^-} > 0.2, \quad \Delta R_{e^\pm\mu} > 0.3$ |

Table 6.1.: Definition of `inclusive setup` and `fiducial setup` [172] [161] used in the phenomenological analysis for $pp \to \mu^+\nu_\mu e^+e^-$ production. Note that leptons are always considered dressed.

In the comparison to data in section 6.4.3, we employ the corresponding Rivet routine [173] available in the `HEPdata` webpage `https://www.hepdata.net/record/ins1720438`. Note that in table 6.1, we define the transverse mass of the $W$ boson as:

$$m_{\text{T,W}} = \sqrt{\left(E_{\text{T},\mu} + E_{\text{T},\nu_\mu}\right)^2 - p^2_{\text{T},\mu\nu_\mu}} \quad \text{with} \quad E^2_{\text{T},x} = m_x^2 + p^2_{\text{T},x}\,. \tag{6.21}$$

In the Rivet analysis, the same definition is adopted but in the massless approximation for dressed leptons ($m_x = 0$ in equation (6.21)).

### 6.4.2. Differential distributions

In this section we analyse in detail the different combination schemes of QCD and EW calculations matched to parton showers presented in section 6.2.3. For simplicity, we consider again only the following decay channel: $pp \to \mu^+\nu_\mu e^+e^-$.

Our plots are organised as follows:

- **Main panel**

  We show the $\text{NNLO}_{\text{QCD}}^{(\text{QCD,QED})_{\text{PS}}}$ result (blue and dashed line), in which the $\text{NNLO}_{\text{QCD}}$ computation is supplemented with both QCD and QED showers. In this prediction, EW effects are pure QED effects coming from the parton shower. Moreover, we present our default multiplicative and additive combinations, namely $\mathrm{NNLO}_{\mathrm{QCD+EW}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}$ (magenta and long-dashed curve) and $\mathrm{NNLO}_{\mathrm{QCDxEW}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}$ (green and solid curve).

- **First ratio panel**

  We present the ratio to the $\mathrm{NNLO}_{\mathrm{QCD}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}$ prediction, in order to study the impact of EW effects. Apart from the $\mathrm{NNLO}_{\mathrm{QCD+EW}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}$ and $\mathrm{NNLO}_{\mathrm{QCDxEW}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}$ combinations, we present also the pure $\mathrm{NNLO}_{\mathrm{QCD}}$+PS calculation, namely $\mathrm{NNLO}_{\mathrm{QCD}}^{(\mathrm{QCD})_{\mathrm{PS}}}$ (red and dash-dotted curve). In this way we can observe the impact of both EW effects coming from the fixed-order computation and QED effects originating from the parton shower.

- **Second ratio panel**

  We show the ratio to the default additive combination $\mathrm{NNLO}_{\mathrm{QCD+EW}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}$. In this panel we include the default multiplicative scheme $\mathrm{NNLO}_{\mathrm{QCDxEW}}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}$ and the multiplicative combination in which the EW K-factor is obtained through a fixed-order computation $\mathrm{NNLO}_{\mathrm{QCD}}^{(\mathrm{QCD})_{\mathrm{PS}}} \times \text{K-NLO}_{\mathrm{EW}}^{(\mathrm{f.o.})}$ (brown and dash-dotted curve).

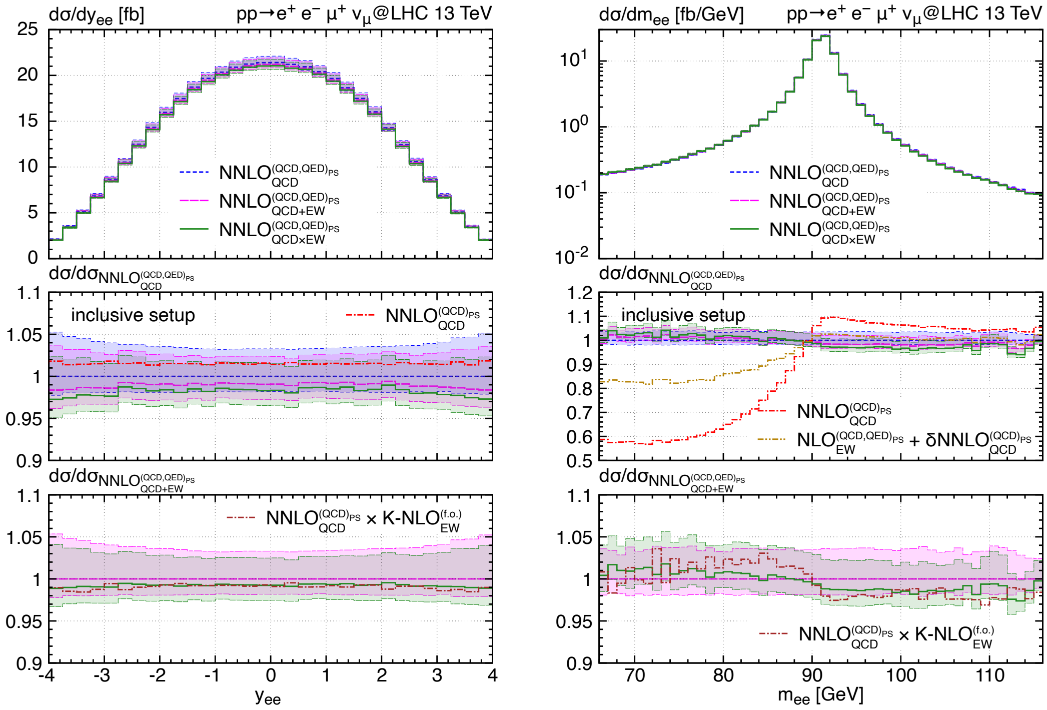


Figure 6.5.: Rapidity $y_{ee}$ (left) and invariant mass $m_{ee}$ (right) of the reconstructed $Z$ boson for $W^+Z$ production in the `inclusive setup`.

We start our discussion with the rapidity $y_{ee}$ of the reconstructed $Z$ boson, shown in the left panel of figure 6.5. EW effects are negative and small (few percent). They are completely covered by scale uncertainties, which have edges of around $+3$–$5\%$ and $-2$–$3\%$ in all predictions. EW corrections are flat in the phase space for this observable: this is not unexpected since $y_{ee}$ is inclusive over QED radiation. For studying the effects of pure QED emissions generated by the shower, we can compare the $\text{NNLO}_{\text{QCD}}^{(\text{QCD})_{\text{PS}}}$ with the $\text{NNLO}_{\text{QCD}}^{(\text{QCD,QED})_{\text{PS}}}$ results: these effects are at the level of $-1$–$2\%$. A comparison with the default combinations with EW calculations shows that effects of pure EW origin are of order $-2$–$3\%$. The default additive and multiplicative combinations ($\text{NNLO}_{\text{QCD+EW}}^{(\text{QCD,QED})_{\text{PS}}}$ and $\text{NNLO}_{\text{QCDxEW}}^{(\text{QCD,QED})_{\text{PS}}}$) agree at the percent level. Moreover, the multiplicative case with the fixed-order K-factor is almost identical to the default multiplicative combination, which means that $y_{ee}$ is weakly affected by secondary photon emissions.

The right panel of figure 6.5 shows the invariant mass $m_{ee}$ of the reconstructed $Z$ boson. In this plot we add one more combination, $\text{NLO}_{\text{EW}}^{(\text{QCD,QED})_{\text{PS}}} + \delta\text{NNLO}_{\text{QCD}}^{(\text{QCD})_{\text{PS}}}$ (beige and dash-double-dotted curve). This combination is the additive scheme in which QED effects generated by the parton shower are not encoded on top of the NNLO QCD computation. For this observable, we expect important distortions when including QED effects, as already pointed out for Drell-Yan production in [174, 175, 176]. In fact, considering the pure QCD computation $\text{NNLO}_{\text{QCD}}^{(\text{QCD})_{\text{PS}}}$ is not suitable to describe the $Z$ resonance, as this calculation does not take into account large QED collinear effects that shift events from above to below the Breit-Wigner peak. These effects are of order 40% at around $m_{ee} = 70$ GeV. A similar result can be observed in the new $\text{NLO}_{\text{EW}}^{(\text{QCD,QED})_{\text{PS}}} + \delta\text{NNLO}_{\text{QCD}}^{(\text{QCD})_{\text{PS}}}$ combination: the distortion of the curve is related to missing large QCD–QED effects (even though beyond accuracy) given by the interplay of the QED shower with the NNLO QCD calculation. Moreover, we note an excellent agreement between $\text{NNLO}_{\text{QCD}}^{(\text{QCD,QED})_{\text{PS}}}$ and the default additive and multiplicative combinations. The multiplicative $\text{NNLO}_{\text{QCDxEW}}^{(\text{QCD,QED})_{\text{PS}}}$ and $\text{NNLO}_{\text{QCD}}^{(\text{QCD})_{\text{PS}}} \times \text{K-NLO}_{\text{EW}}^{(\text{f.o.})}$combinations are in good agreement (1–2% level), which validates the matching of the EW computation with the QED shower in a resonance-aware way, as done in the POWHEG-BOX-RES framework.

We now turn to high-energy tails of relevant distributions, showing results in both `inclusive setup` and `fiducial setup` as defined in table 6.1. Notice that we adopt a logarithmic scale and binning on the $x$-axis. In the following plots, we add one more combination in the second ratio inset, namely $\text{NNLO}_{\text{QCD}}^{(\text{QCD,QED})_{\text{PS}}} \times \text{K-NLO}_{\text{EW}}^{(\text{QED})_{\text{PS}}}$ (orange and dash-double-dotted curve), in which the QCD shower is not applied to the EW K-factor.

Figure 6.6 shows the missing transverse momentum $p_{\text{T,miss}}$ in the `inclusive setup` (left panel) and `fiducial setup` (right panel). The two setups show similar results. Pure QED effects are small (at the percent level), as observed comparing the $\text{NNLO}_{\text{QCD}}^{(\text{QCD})_{\text{PS}}}$ and $\text{NNLO}_{\text{QCD}}^{(\text{QCD,QED})_{\text{PS}}}$ curves. By contrast, EW contributions are strongly enhanced in the deep tails due to the appearance of EW Sudakov logarithms. As clearly shown in the second ratio panel, there is a large discrepancy between the default additive and multiplicative schemes, and this effect is originated by large NLO QCD corrections that give rise to gi-

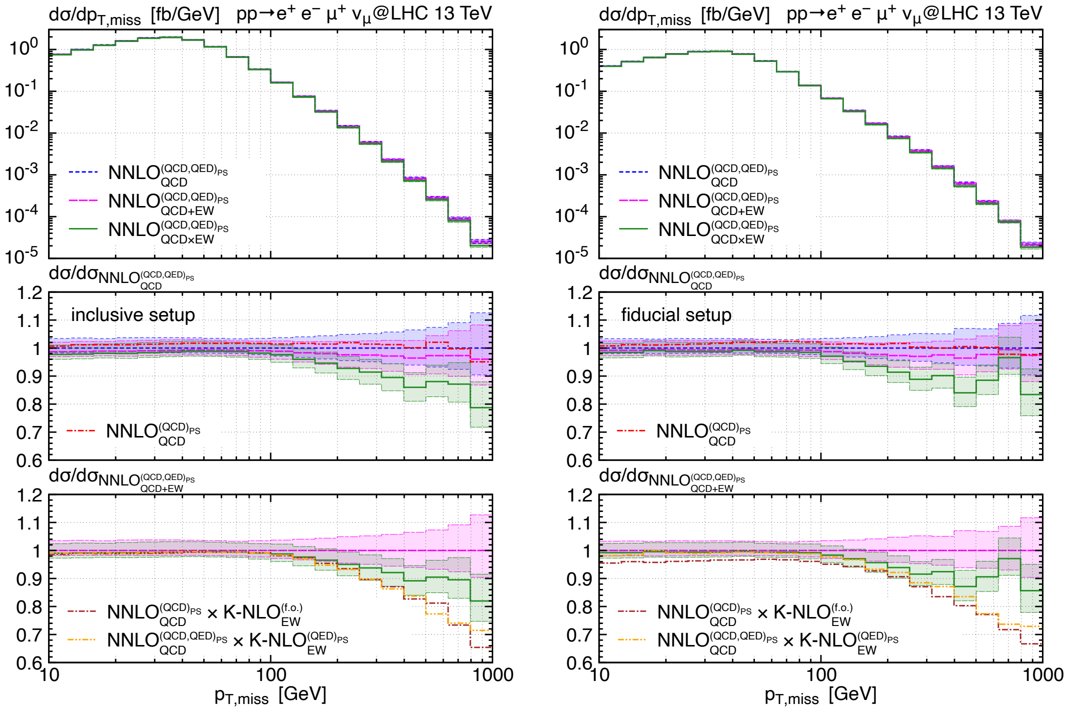


Figure 6.6.: Missing transverse momentum $p_{\mathrm{T,miss}}$ for $W^+Z$ production in the `inclusive setup` (left) and `fiducial setup` (right).

ant K-factors. Moreover, the default multiplicative combination does not agree with the $\mathrm{NNLO}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}_{\mathrm{QCD}} \times \mathrm{K\text{-}NLO}^{(\mathrm{QED})_{\mathrm{PS}}}_{\mathrm{EW}}$ and the $\mathrm{NNLO}^{(\mathrm{QCD})_{\mathrm{PS}}}_{\mathrm{QCD}} \times \mathrm{K\text{-}NLO}^{(\mathrm{f.o.})}_{\mathrm{EW}}$ predictions: this discrepancy arises from giant K-factors that are related to QCD emissions generated by the shower. We explicitly verified that, applying a suitable dynamic veto that suppresses hard-$V$+jet topologies, the three multiplicative combinations are in perfect agreement.

In figures 6.7 and 6.8 we show the transverse momenta of the charged leading lepton $p_{\mathrm{T},\ell_1}$ and the charged subleading lepton $p_{\mathrm{T},\ell_2}$. We note a similar behaviour to the $p_{\mathrm{T,miss}}$ distribution. More precisely, EW effects are negative and strongly enhanced in the high-energy tails. The observed discrepancy among the different multiplicative combinations is related to giant K-factors, as explained for the $p_{\mathrm{T,miss}}$ observable. Note that this discrepancy is absent for the subleading lepton, as this observable is less affected by giant K-factors. QED effects have a non-trivial impact at low transverse momentum, as observed comparing $\mathrm{NNLO}^{(\mathrm{QCD})_{\mathrm{PS}}}_{\mathrm{QCD}}$ with $\mathrm{NNLO}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}_{\mathrm{QCD}}$ in the first ratio panel and comparing $\mathrm{NNLO}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}_{\mathrm{QCDxEW}}$ and $\mathrm{NNLO}^{(\mathrm{QCD})_{\mathrm{PS}}}_{\mathrm{QCD}} \times \mathrm{K\text{-}NLO}^{(\mathrm{f.o.})}_{\mathrm{EW}}$ in the second ratio panel.

Figure 6.9 shows the invariant mass of the three charged leptons $m_{3\ell}$. Also for this observable, EW corrections are negative and visible in the deep tails of the distribution. When comparing the `inclusive setup` (left panel) to the `fiducial setup` (right panel), we observe a significant enhancement of the EW effects. For example, at around $m_{3\ell} \sim 2\,\mathrm{TeV}$, EW

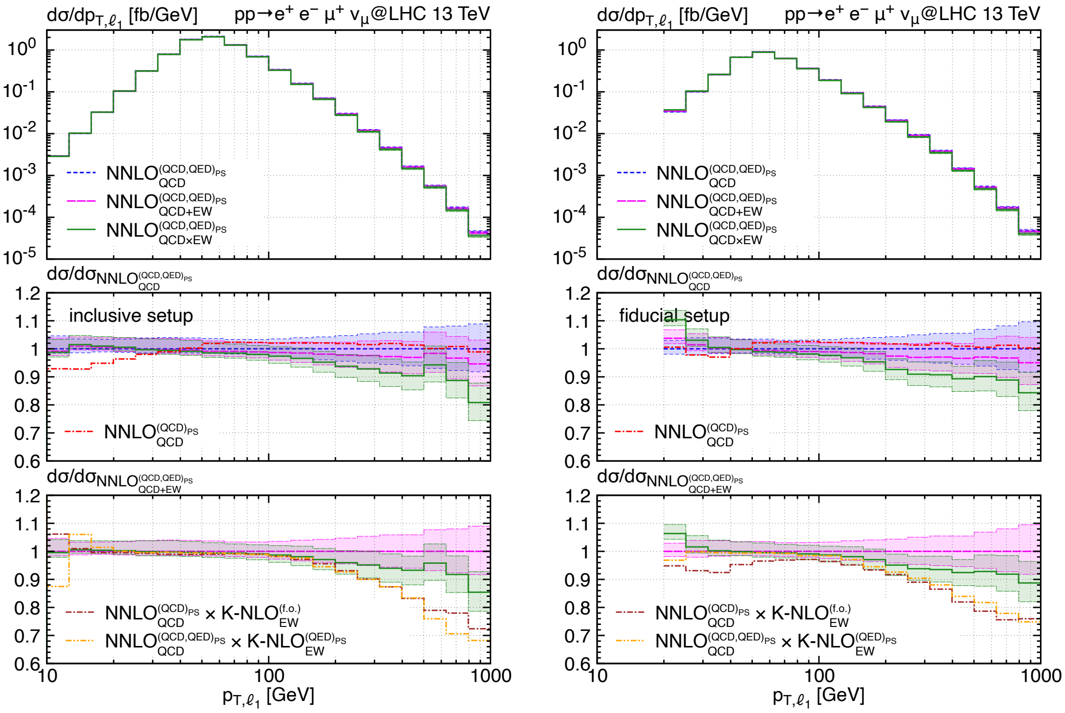


Figure 6.7.: Transverse momentum of the charged leading lepton $p_{\mathrm{T},\ell_1}$ for $W^+Z$ production in the `inclusive setup` (left) and `fiducial setup` (right).

corrections are about $-10\%$ in the `inclusive setup` while they reach about $-20$–$30\%$ in the `fiducial setup`. This effect is due to the definition of the fiducial phase space. In fact, for large $m_{3\ell}$ values leptons are produced at very large rapidities in the `inclusive setup` and, in these topologies, not all the Mandelstam invariants $s_{ij}$ are large. These invariants appear in the EW (double) Sudakov logarithms ($\ln^2(|s_{ij}|/m_W^2)$), which thus result suppressed. By contrast, when applying fiducial cuts in the `fiducial setup`, lepton rapidities are constrained and the very forward regions are excluded from the definition of the phase space, thus leading to the usual enhancement of EW effects in the deep tails.

In all the plots presented in this section, we did not show results associated with the $\mathrm{NNLO}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}_{\mathrm{QCD}} + \delta\mathrm{NLO}^{(\mathrm{QED})_{\mathrm{PS}}}_{\mathrm{EW}}$ and the $\mathrm{NLO}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}_{\mathrm{EW}} \times \text{K-NNLO}^{(\mathrm{QCD})_{\mathrm{PS}}}_{\mathrm{QCD}}$ predictions, even though they were listed as possible combination schemes in section 6.2.3. The reason is that we did not find any significant discrepancy with the $\mathrm{NNLO}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}_{\mathrm{QCD+EW}}$ and the $\mathrm{NNLO}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}_{\mathrm{QCDxEW}}$ calculations, so we refrained from showing results with these combinations. We stress that they should be considered equally appropriate and valid.

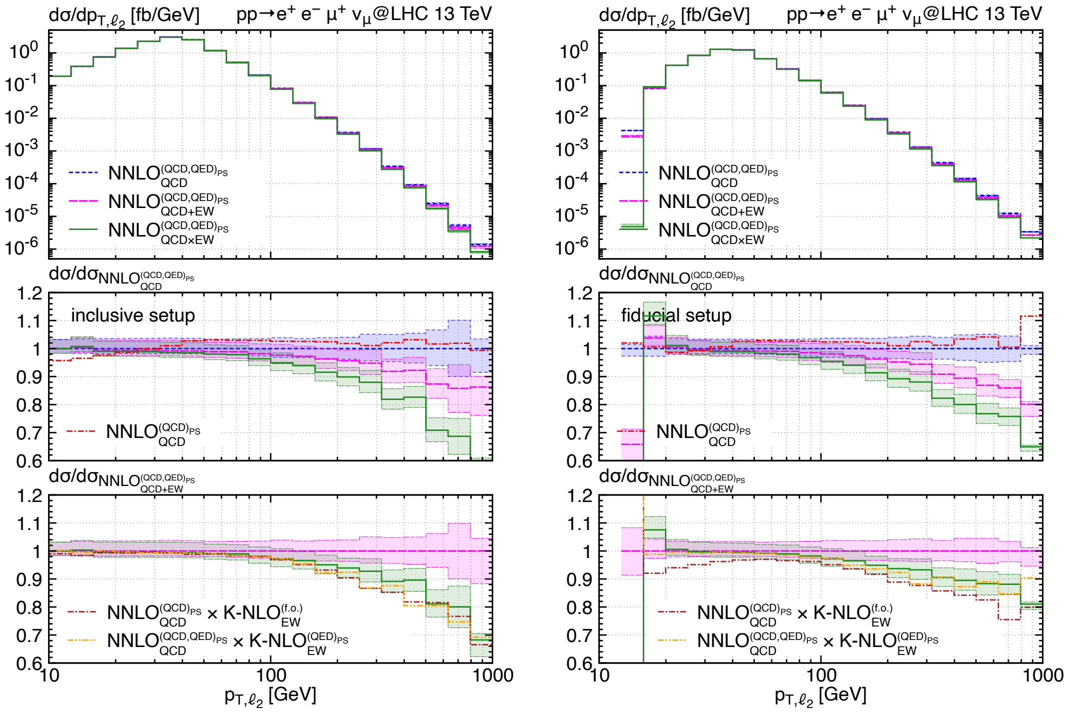


Figure 6.8.: Transverse momentum of the charged subleading lepton $p_{\mathrm{T},\ell_2}$ for $W^+Z$ production in the `inclusive setup` (left) and `fiducial setup` (right).

### 6.4.3. Comparison to data

For the comparison to data, we consider the $\mathrm{NNLO}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}_{\mathrm{QCDxEW}}$ prediction as our default result. ATLAS data are taken from [161] and we adopt the corresponding RIVET analysis available in `https://www.hepdata.net/record/ins1720438`. The fiducial cuts are given by the `fiducial setup` described in table 6.1. In this section, the results are the average of all the possible decay channels of $W^\pm Z$ production ($e^+e^-e^\pm\nu_e$, $\mu^+\mu^-\mu^\pm\nu_\mu$, $e^+e^-\mu^\pm\nu_\mu$, $e^+e^-\mu^\pm\nu_\mu$).

Figure 6.10 shows a comparison of our default result with MPI effects (blue and solid curve), without MPI effects (red and dashed curve) and ATLAS data. In our results, we include effects from hadronization. Note that we show the absolute cross section per bin in all the plots. The last bin of unbound distributions (indicated with ∞) should be seen as an overflow bin. We consider the transverse momenta of the $Z$ and $W$ bosons, $p_{\mathrm{T},Z}$ and $p_{\mathrm{T},W}$ respectively, and of the neutrino $p_{\mathrm{T},\nu}$. We also show the transverse mass of the colour singlet $m_{\mathrm{T},WZ}$ defined as

$$m_{\mathrm{T},WZ} = \sqrt{\left(\sum_{i=1}^{4} p_{\mathrm{T},i}\right)^2 - p^2_{\mathrm{T},WZ}}\,, \tag{6.22}$$

where $i$ runs over the three charged leptons and the neutrino. Moreover, we show the

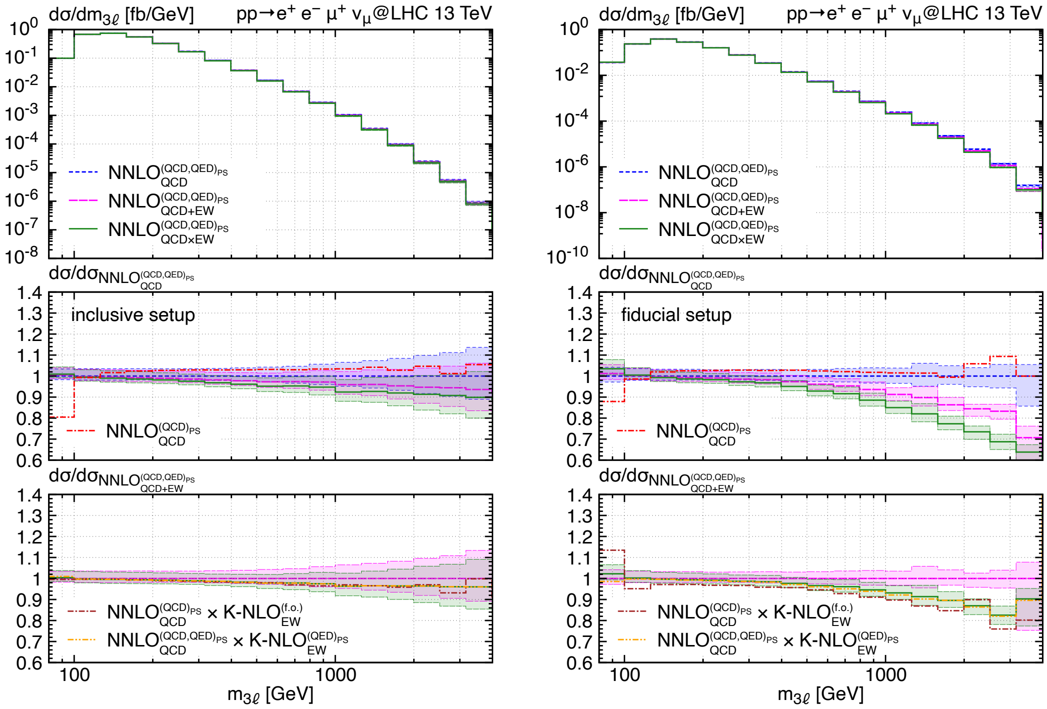


Figure 6.9.: Invariant mass of the three charged leptons for $W^+Z$ production in the `inclusive setup` (left) and `fiducial setup` (right).

difference in the azimuthal angle between the two vector bosons $\Delta\phi_{WZ}$ and the absolute rapidity difference between the $Z$ boson and the charged lepton coming from the $W$ $|y_Z - y_{\ell_W}|$.

Our default MiNNLO$_{\text{PS}}$ prediction is in excellent agreement with data, both in the bulk of the cross section where QCD is dominant, and in the tails of distributions where EW effects become large. Note that all the shown distributions are genuinely NNLO QCD and NLO EW accurate, except for $\Delta\phi_{WZ}$. In fact, the latter is equal to $\pi$ at LO and the entire spectrum is filled starting from NLO calculations, as a recoil of the $WZ$ system on extra radiation is required. This observable is formally only NLO QCD accurate, which is reflected in the slightly larger error bar at low $\Delta\phi_{WZ}$. MPI effects lower the MiNNLO$_{\text{PS}}$ prediction by 5%, making the agreement with data slightly worse. We notice that theoretical predictions are extremely precise, as scale variation leads to a few percent error bars. By contrast, data have much larger uncertainties. Nevertheless, these measurements will improve significantly when Run III data is considered in the future, thus performing again this comparison will be particularly relevant.

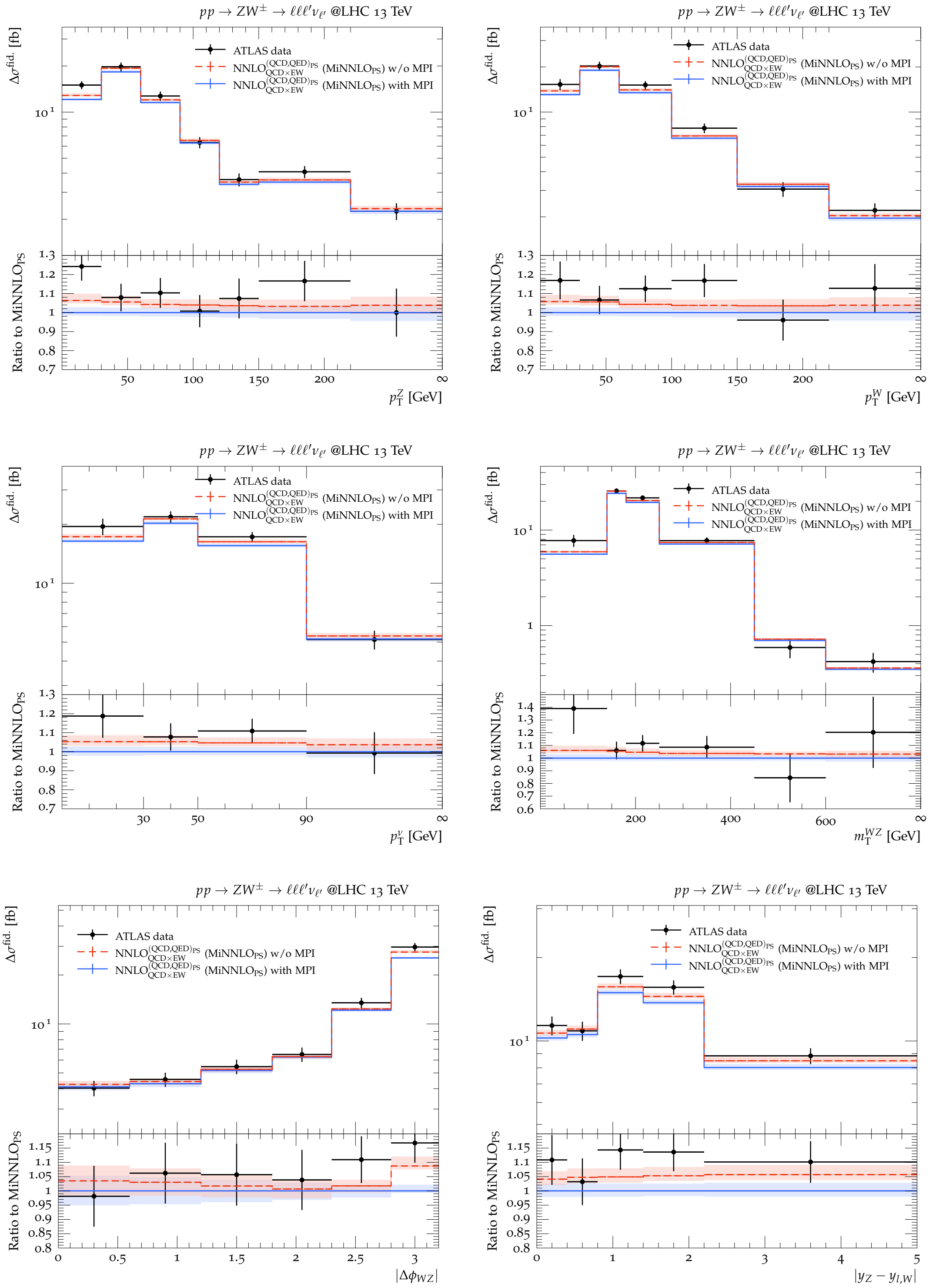


Figure 6.10.: Comparison of the $\mathrm{NNLO}^{(\mathrm{QCD,QED})_{\mathrm{PS}}}_{\mathrm{QCDxEW}}$ result with MPI effects (blue and solid curve), without MPI effects (red and dashed curve) and ATLAS data [161].

# Part III.
# CONCLUSIONS

# 7. Conclusions

High-precision collider phenomenology is the common thread of this thesis. Since the Standard Model cannot be the final fundamental theory describing the subatomic world and no clear hints of new physics has been observed so far, precision physics has become an indispensable element for any possible progress in LHC physics. New physics effects can appear as small deviations when comparing data to theoretical calculations, which directly translates into the necessity of obtaining both data and theoretical predictions at ultimate precision.

In recent years the LHC has entered the so-called precision era: experimental measurements are reaching an unprecedented level of precision in many hadron-collider observables especially with the upcoming improvement to the High Luminosity phase. In order to interpret correctly these data, theoretical calculations should match the same level of precision. Moreover, a direct data/theory comparison can be performed only through realistic and flexible theoretical simulations. In this thesis we analysed in detail one of the main building blocks for obtaining precise and realistic theoretical predictions, namely the matching between fixed-order calculations and parton showers.

Fixed-order calculations and parton shower simulations are at the core of the description of a collider event in the perturbative regime. These two tools are complementary approaches that can be fully exploited to obtain a realistic and accurate event simulation when consistently combined. In particular, a fixed-order computation is suitable for a precise description of the hard scattering process, but it does not provide a full simulation of the collider event. By contrast, a parton shower simulation describes the realistic evolution of a collider event from the high-energy limit down to the detector level in the soft and collinear approximation but only at low accuracy. Matching the two descriptions, while keeping the best features of both, thus becomes crucial for high-precision physics.

Nowadays, this combination can be performed at NNLO+PS accuracy, which means keeping the NNLO accuracy of the fixed-order prediction without spoiling the structure of the parton shower. Among the methods available in the literature, the MiNNLO$_{\text{PS}}$ framework plays an important role because it is a flexible and numerically efficient tool that requires neither a CPU-demanding a-posteriori reweighting nor the introduction of an unphysical merging scale for partitioning the phase space according to the jet multiplicity. The main goal of this thesis is to present some cutting-edge results obtained using the MiNNLO$_{\text{PS}}$ method for key processes at the LHC.

Part I presents a review of the theoretical ingredients that are needed to construct a NNLO+PS accurate event generator. We start with a detailed overview of both fixed-order calculations and parton shower simulations in chapter 2, highlighting their strengths and weaknesses and showing why they are complementary tools that can be combined consistently.

The problem of the matching at NNLO+PS accuracy is presented in chapter 3: reaching NNLO+PS accuracy is highly non-trivial and it first requires a deep understanding of the matching procedure in a simpler case, namely NLO+PS. We thus present the basics of the POWHEG method for the generation of NLO+PS accurate events in section 3.1. The improvement of this framework into the MiNLO and the MiNLO′ methods is described in section 3.2 where a merging of different jet multiplicities is performed without the introduction of unphysical slicing parameters. The MiNNLO$_{\text{PS}}$ method is finally presented in section 3.3.

The core of this thesis is Part II, where we present phenomenological results for relevant processes at the LHC.

Chapter 4 is dedicated to precision Higgs physics and jet flavours. We discuss the implementation of a MiNNLO$_{\text{PS}}$-based event generator within the POWHEG-BOX-RES and POWHEG-BOX-V2 frameworks for the associated production of a Higgs and a vector boson ($pp \to VH$, with $V = W^{\pm}, Z$) with subsequent decay of the Higgs boson to a pair of bottom quarks ($H \to b\bar{b}$). The vector boson $V$ can decay to any possible leptonic final state. This process is crucial for precision measurements in the Higgs sector as it yields the highest sensitivity to the determination of the bottom Yukawa coupling. The matching with the parton shower (PYTHIA8) is non-trivial because the starting scales of the emissions generated by the shower are different in the production and decay stages of the Higgs boson. In our phenomenological results for 13TeV LHC collisions, we compare our MiNNLO$_{\text{PS}}$ results, MiNLO′ results and predictions obtained using PYTHIA8 for generating the decay of the Higgs boson (formally, this is a LO decay). The outcome of our study is that the inclusion of NNLO corrections is fundamental for improving our description of the process at hand, as it induces a ∼5% increase in the normalization on top of the MiNLO′ predictions and it determines an important reduction of theoretical uncertainties. Moreover, employing PYTHIA8 for generating the decay of the Higgs boson is not suitable for correctly describing certain observables (e.g. the absolute difference in rapidity of the two reconstructed $b$-quarks is highly distorted in the very forward region). We presented also a detailed analysis of the impact of the jet-clustering algorithm that is used for the identification of $b$-jets. We performed a comparison among the flavour-$k_T$ clustering algorithm, which is needed for an infrared safe definition of jet flavour, and usual clustering algorithms that are insensitive to the flavour of particles, namely the $k_T$ and anti-$k_T$ algorithms. Using the flavour-$k_T$ algorithm leads to significant distortions of relevant distributions in certain kinematic regions of the phase space, which led us to the conclusion that this algorithm should be used only when it is applied also on the experimental side. This is usually not the case, as it is still too challenging for a practical application. More progress is surely needed for defining a flavour of jets that is both theoretically consistent and applicable to experimental measurements.

Chapter 5 presents an extension of the implementation of $ZH$ production with $H \to b\bar{b}$ decay to the SMEFT. We include in our computation a subset of dimension-six operators that are relevant up to NNLO accuracy in QCD for the process at hand. We thus follow the strategy presented for the same calculation in the SM (in chapter 4) for implementing a Monte Carlo event generator at NNLO+PS accuracy in the SMEFT. While for the production stage of the Higgs boson the SMEFT effects we are considering are completely factorized, in the $H \to b\bar{b}$

decay both factorizable and non-factorizable corrections appear. In the non-factorizable contributions we encode corrections that are formally $N^3$LO accurate, due to the weak constraints on the corresponding Wilson coefficients. We study two main scenarios where we separately switch on factorizable and non-factorizable effects. This is done considering all coefficients equal to zero but one, the value of which is set according to its current experimental bounds. Factorizable corrections determine only a shift in the normalisation with respect to the SM, while non-factorizable contributions are more interesting because they alter the shape of relevant distributions. In particular, we identified two kinematic observables that are suitable for increasing our sensitivity to the considered SMEFT effects. The first one is the invariant mass of the $b\bar{b}$ system, which is reconstructed as the pair of $b$-jets whose invariant mass is closest to the Higgs mass; the second distribution is the invariant mass of the $b\bar{b}j$ system, which is reconstructed as the set of two $b$-jets and one additional jet whose three-jet invariant mass is closest to the Higgs-boson mass. These two distributions receive large SMEFT effects in the low-mass region. Moreover, these effects are dependent on the exact jet definition because they change when varying the jet radius used in the clustering procedure. This work is the starting point of a more challenging project in which NNLO+PS accuracy in the SMEFT is reached for an enlarged set of operators, including e.g. non-trivial insertions of pure EW operators.

In chapter 6 we present a strategy for targeting NNLO QCD and NLO EW accuracy matched to parton showers, namely $\text{NNLO}_{\text{QCD}}$+PS and $\text{NLO}_{\text{EW}}$+PS accuracy. The combination of QCD and EW corrections matched to PS is done a posteriori, at the level of differential distributions. We present phenomenological results for $W^{\pm}Z$ production, but we stress that this strategy is fully general. $W^{\pm}Z$ production represents a precision test for the SM because it provides direct access to triple gauge couplings, thus probing the gauge structure of the EW sector. We implemented a $\text{MiNNLO}_{\text{PS}}$-based event generator for reaching $\text{NNLO}_{\text{QCD}}$+PS accuracy within the POWHEG-BOX-RES framework. Moreover, we implemented a POWHEG code for producing $\text{NLO}_{\text{EW}}$+PS accurate events, again within POWHEG-BOX-RES. The matching with the parton shower is not trivial because QCD and QED emissions should respect different constraints according to the emissions already present at Les Houches Event level. We propose seven different combination schemes that differ only by higher-order corrections. NNLO QCD computations matched with both QCD and QED showers provide a good description of the bulk of the cross section, while EW effects become increasingly dominant in the tails of kinematic distributions. This project is a first step towards the possibility of implementing a single MC event generator that reaches the targeted accuracy in one calculation. Concretely, we would like to extend the $\text{MiNNLO}_{\text{PS}}$ method by including NLO EW effects. This extension is extremely challenging as it requires a deep understanding of the IR structure of mixed QCD-EW contributions in the $\text{MiNNLO}_{\text{PS}}$ formalism.

The results presented in this thesis are the state of the art for precision phenomenology. Reaching NNLO+PS accuracy is nowadays mandatory for a correct understanding of LHC data and the $\text{MiNNLO}_{\text{PS}}$ method has revealed to be a powerful and flexible tool for targeting this accuracy. In view of the imminent High Luminosity upgrade, we believe that our results and developed tools will be very valuable for the high-energy community, especially for current and future experimental analyses. In light of this, our MC event generators will be made publicly available on the POWHEG-BOX webpage.

# APPENDIX

# A. Differential distributions for $W^-H$ production with $H \to b\bar{b}$ decay at NNLO+PS in the SM

In this section we present differential distributions for $pp \to W^-H \to e^-\bar{\nu}_e b\bar{b}$ analogous to the results presented in section 4.4.3. We refrain from commenting these results further as the same conclusions that were discussed for $W^+H$ production still hold.

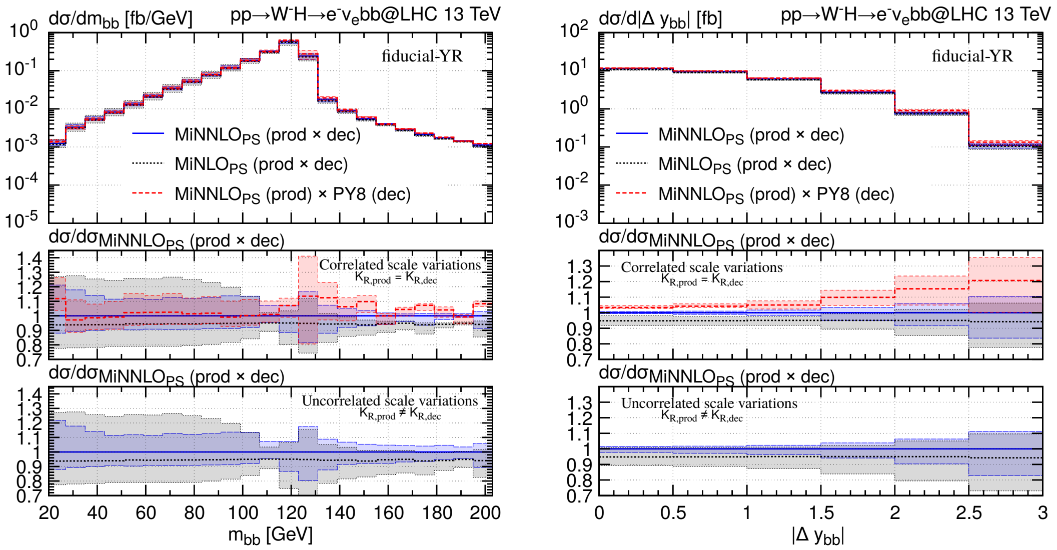


Figure A.1.: Differential distributions for $e^-\bar{\nu}_e b\bar{b}$ production with `fiducial-YR` cuts (table 4.5). See text for more details.

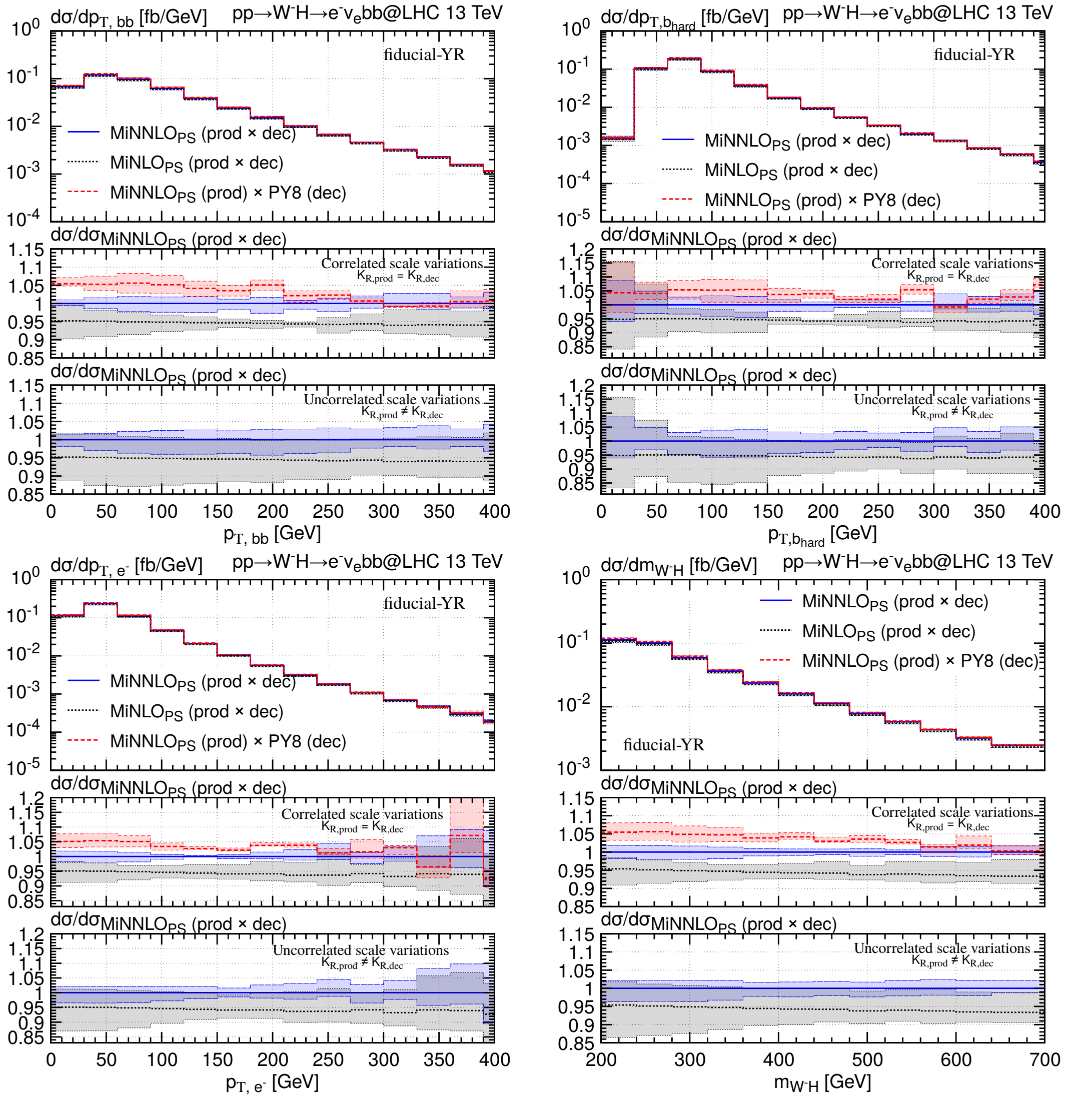


Figure A.2.: Differential distributions for $e^-\bar{\nu}_e b\bar{b}$ production with `fiducial-YR` cuts (table 4.5). See text for more details.

# Acknowledgments

First of all, I would like to thank my supervisor Giulia Zanderighi for the constant support during my doctoral journey. Thank you for teaching me so much, I will treasure this experience. A big thank you to all the members of my research group for creating a wonderful work environment.

I would like to express my gratitude to my entire family for their continuous support. I know it hasn't been easy being apart, but you have always embraced my choices. And a huge thank you to Daniele, who has always believed in me from the beginning: without him, everything would have been different.

I thank my dearest friends Chiara and Anita, because they are always with me. The same goes for Carlo, Davide, Giorgia and Alessandro: you are always in my thoughts.

Now, I want to thank all the people who have been part of my daily life here in Munich and have made this experience truly unique. I thank Alessandro and Christian for being the best office mates, and Thomas for being a fellow office mate even though he is technically based in another office. You have been like a small family to me, and I will miss you immensely. Thanks to my favourite postdoc Giovanni, because even though we never fully understand each other, we always get along. Special thanks to Ana, who always knows how to lift your mood (and for the rabbit), and to Cesare and Checco, for all the wonderful evenings together. Giordano and Giacomo, thank you for the best coffee breaks. A huge thank you to all my fellow PhD students at MPP: I can't mention all of you, but you have been invaluable.

# Bibliography

[1] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization for Short Distance Hadron - Hadron Scattering*, *Nucl. Phys. B* **261** (1985) 104–142.

[2] S. Zanoli, M. Chiesa, E. Re, M. Wiesemann and G. Zanderighi, *Next-to-next-to-leading order event generation for VH production with $H \to b\bar{b}$ decay*, *JHEP* **07** (2022) 008, [`2112.04168`].

[3] U. Haisch, D. J. Scott, M. Wiesemann, G. Zanderighi and S. Zanoli, *NNLO event generation for $pp \to Zh \to \ell^+\ell^- b\bar{b}$ production in the SM effective field theory*, *JHEP* **07** (2022) 054, [`2204.00663`].

[4] J. M. Lindert, D. Lombardi, M. Wiesemann, G. Zanderighi and S. Zanoli, *WZ production at NNLO QCD and NLO EW matched to parton showers with MiNNLO$_{PS}$*, *JHEP* **11** (2022) 036, [`2208.12660`].

[5] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070, [`0709.2092`].

[6] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.

[7] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and collider physics*, *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **8** (1996) 1–435.

[8] 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`].

[9] PARTICLE DATA GROUP collaboration, P. A. Zyla et al., *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01.

[10] T. Kinoshita, *Mass singularities of Feynman amplitudes*, *J. Math. Phys.* **3** (1962) 650–677.

[11] T. D. Lee and M. Nauenberg, *Degenerate Systems and Mass Singularities*, *Phys. Rev.* **133** (1964) B1549–B1562.

[12] G. Luisoni and S. Marzani, *QCD resummation for hadronic final states*, *J. Phys. G* **42** (2015) 103101, [`1505.04084`].

[13] PARTICLE DATA GROUP collaboration, R. L. Workman and Others, *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.

[14] 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) 641–653.

[15] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438–450.

[16] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys.* **B126** (1977) 298–318.

[17] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and Collider Physics*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 1996, 10.1017/CBO9780511628788.

[18] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, [`hep-ph/0409146`].

[19] S. Catani and M. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, *Nucl. Phys.* **B485** (1997) 291–419, [`hep-ph/9605323`].

[20] 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.* **B627** (2002) 189–265, [`hep-ph/0201036`].

[21] S. Frixione, Z. Kunszt and A. Signer, *Three jet cross-sections to next-to-leading order*, *Nucl. Phys. B* **467** (1996) 399–442, [`hep-ph/9512328`].

[22] S. Frixione, *A General approach to jet cross-sections in QCD*, *Nucl. Phys. B* **507** (1997) 295–314, [`hep-ph/9706545`].

[23] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, A. Huss and T. A. Morgan, *Precise QCD predictions for the production of a Z boson in association with a hadronic jet*, *Phys. Rev. Lett.* **117** (2016) 022001, [`1507.02850`].

[24] M. Czakon, *A novel subtraction scheme for double-real radiation at NNLO*, *Phys. Lett.* **B693** (2010) 259–268, [`1005.0274`].

[25] M. Cacciari, F. A. Dreyer, A. Karlberg, G. P. Salam and G. Zanderighi, *Fully Differential Vector-Boson-Fusion Higgs Production at Next-to-Next-to-Leading Order*, *Phys. Rev. Lett.* **115** (2015) 082002, [`1506.02660`].

[26] F. Caola, K. Melnikov and R. Röntsch, *Nested soft-collinear subtractions in NNLO QCD computations*, *Eur. Phys. J. C* **77** (2017) 248, [`1702.01352`].

[27] G. Somogyi, Z. Trocsanyi and V. Del Duca, *Matching of singly- and doubly-unresolved limits of tree-level QCD squared matrix elements*, *JHEP* **06** (2005) 024, [`hep-ph/0502226`].

[28] G. Bertolotti, L. Magnea, G. Pelliccioli, A. Ratti, C. Signorile-Signorile, P. Torrielli and S. Uccirati, *NNLO subtraction for any massless final state: a complete analytic expression*, `2212.11190`.

[29] J. C. Collins, D. E. Soper and G. F. Sterman, *Transverse Momentum Distribution in Drell-Yan Pair and W and Z Boson Production*, *Nucl. Phys. B* **250** (1985) 199–224.

[30] S. Catani, L. Trentadue, G. Turnock and B. R. Webber, *Resummation of large logarithms in e+ e- event shape distributions*, *Nucl. Phys. B* **407** (1993) 3–42.

[31] R. Bonciani, S. Catani, M. L. Mangano and P. Nason, *Sudakov resummation of multiparton QCD cross-sections*, *Phys. Lett. B* **575** (2003) 268–278, [`hep-ph/0307035`].

[32] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna and E. Norrbin, *High-energy physics event generation with PYTHIA 6.1*, *Comput. Phys. Commun.* **135** (2001) 238–259, [`hep-ph/0010017`].

[33] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour and B. R. Webber, *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, *JHEP* **01** (2001) 010, [`hep-ph/0011363`].

[34] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert and J. Winter, *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007, [`0811.4622`].

[35] A. Buckley et al., *General-purpose event generators for LHC physics*, *Phys. Rept.* **504** 145–233, [`1101.2599`].

[36] T. Sjostrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026, [`hep-ph/0603175`].

[37] S. Catani, B. R. Webber and G. Marchesini, *QCD coherent branching and semiinclusive processes at large x*, *Nucl. Phys. B* **349** (1991) 635–654.

[38] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029, [`hep-ph/0204244`].

[39] K. Hamilton, P. Nason and G. Zanderighi, *MINLO: Multi-Scale Improved NLO*, *JHEP* **10** (2012) 155, [`1206.3572`].

[40] K. Hamilton, P. Nason, C. Oleari and G. Zanderighi, *Merging H/W/Z + 0 and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching*, *JHEP* **05** (2013) 082, [`1212.4504`].

[41] 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`].

[42] N. Lavesson and L. Lonnblad, *Extending CKKW-merging to One-Loop Matrix Elements*, *JHEP* **12** (2008) 070, [`0811.2912`].

[43] S. Alioli, K. Hamilton and E. Re, *Practical improvements and merging of POWHEG simulations for vector boson production*, *JHEP* **09** (2011) 104, [`1108.0909`].

[44] S. Hoeche, F. Krauss, M. Schonherr and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, *JHEP* **04** (2013) 027, [`1207.5030`].

[45] T. Gehrmann, S. Hoche, F. Krauss, M. Schonherr and F. Siegert, *NLO QCD matrix elements + parton showers in* $e^+e^-$ *—> hadrons*, *JHEP* **01** (2013) 144, [`1207.5031`].

[46] R. Frederix and S. Frixione, *Merging meets matching in MC@NLO*, *JHEP* **12** (2012) 061, [`1209.6215`].

[47] S. Alioli, C. W. Bauer, C. J. Berggren, A. Hornig, F. J. Tackmann, C. K. Vermilion, J. R. Walsh and S. Zuberi, *Combining Higher-Order Resummation with Multiple NLO Calculations and Parton Showers in GENEVA*, *JHEP* **09** (2013) 120, [`1211.7049`].

[48] S. Plätzer, *Controlling inclusive cross sections in parton shower + matrix element merging*, *JHEP* **08** (2013) 114, [`1211.5467`].

[49] L. Lönnblad and S. Prestel, *Merging Multi-leg NLO Matrix Elements with Parton Showers*, *JHEP* **03** (2013) 166, [`1211.7278`].

[50] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *QCD matrix elements + parton showers*, *JHEP* **11** (2001) 063, [`hep-ph/0109231`].

[51] D. de Florian and M. Grazzini, *The Structure of large logarithmic corrections at small transverse momentum in hadronic collisions*, *Nucl. Phys. B* **616** (2001) 247–285, [`hep-ph/0108273`].

[52] T. Becher and M. Neubert, *Drell-Yan Production at Small* $q_T$*, Transverse Parton Distributions and the Collinear Anomaly*, *Eur. Phys. J. C* **71** (2011) 1665, [`1007.4005`].

[53] C. Davies and W. Stirling, *Nonleading Corrections to the Drell-Yan Cross-Section at Small Transverse Momentum*, *Nucl. Phys. B* **244** (1984) 337–348.

[54] D. de Florian and M. Grazzini, *Next-to-next-to-leading logarithmic corrections at small transverse momentum in hadronic collisions*, *Phys. Rev. Lett.* **85** (2000) 4678–4681, [`hep-ph/0008152`].

[55] C. T. H. Davies, B. R. Webber and W. J. Stirling, *Drell-Yan Cross-Sections at Small Transverse Momentum*, .

[56] K. Hamilton, P. Nason, E. Re and G. Zanderighi, *NNLOPS simulation of Higgs boson production*, *JHEP* **10** (2013) 222, [`1309.0017`].

[57] A. Karlberg, E. Re and G. Zanderighi, *NNLOPS accurate Drell-Yan production*, *JHEP* **09** (2014) 134, [`1407.2940`].

[58] W. Astill, W. Bizon, E. Re and G. Zanderighi, *NNLOPS accurate associated HW production*, *JHEP* **06** (2016) 154, [`1603.01620`].

[59] W. Astill, W. Bizoń, E. Re and G. Zanderighi, *NNLOPS accurate associated HZ production with $H \to b\bar{b}$ decay at NLO*, *JHEP* **11** (2018) 157, [`1804.08141`].

[60] E. Re, M. Wiesemann and G. Zanderighi, *NNLOPS accurate predictions for $W^+W^-$ production*, *JHEP* **12** (2018) 121, [`1805.09857`].

[61] W. Bizoń, E. Re and G. Zanderighi, *NNLOPS description of the $H \to b\bar{b}$ decay with MiNLO*, *JHEP* **06** (2020) 006, [`1912.09982`].

[62] S. Alioli, C. W. Bauer, C. Berggren, F. J. Tackmann, J. R. Walsh and S. Zuberi, *Matching Fully Differential NNLO Calculations and Parton Showers*, *JHEP* **06** (2014) 089, [`1311.0286`].

[63] S. Höche, Y. Li and S. Prestel, *Drell-Yan lepton pair production at NNLO QCD with parton showers*, *Phys. Rev. D* **91** (2015) 074015, [`1405.3607`].

[64] S. Höche, Y. Li and S. Prestel, *Higgs-boson production through gluon fusion at NNLO QCD with parton showers*, *Phys. Rev.* **D90** (2014) 054011, [`1407.3773`].

[65] 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`].

[66] D. Lombardi, M. Wiesemann and G. Zanderighi, *Advancing MiNNLO$_{\rm PS}$ to diboson processes: $Z\gamma$ production at NNLO+PS*, *JHEP* **06** (2021) , [`2010.10478`].

[67] J. Mazzitelli, P. F. Monni, P. Nason, E. Re, M. Wiesemann and G. Zanderighi, *Next-to-Next-to-Leading Order Event Generation for Top-Quark Pair Production*, *Phys. Rev. Lett.* **127** (2021) 062001, [`2012.14267`].

[68] D. Lombardi, M. Wiesemann and G. Zanderighi, *Anomalous couplings in $Z\gamma$ events at NNLO+PS and improving $\nu\nu\,^-\gamma$ backgrounds in dark-matter searches*, *Phys. Lett. B* **824** (2022) 136846, [`2108.11315`].

[69] D. Lombardi, M. Wiesemann and G. Zanderighi, *$W^+W^-$ production at NNLO+PS with MINNLO$_{PS}$*, *JHEP* **11** (2021) 230, [`2103.12077`].

[70] L. Buonocore, G. Koole, D. Lombardi, L. Rottoli, M. Wiesemann and G. Zanderighi, *ZZ production at nNNLO+PS with MiNNLO$_{PS}$*, *JHEP* **01** (2022) 072, [`2108.05337`].

[71] J. Mazzitelli, P. F. Monni, P. Nason, E. Re, M. Wiesemann and G. Zanderighi, *Top-pair production at the LHC with MINNLO$_{PS}$*, *JHEP* **04** (2022) 079, [`2112.12135`].

[72] J. Mazzitelli, A. Ratti, M. Wiesemann and G. Zanderighi, *B-hadron production at the LHC from bottom-quark pair production at NNLO+PS*, `2302.01645`.

[73] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr and G. Watt, *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J.* **C75** (2015) 132, [`1412.7420`].

[74] G. P. Salam and J. Rojo, *A Higher Order Perturbative Parton Evolution Toolkit (HOPPET)*, *Comput. Phys. Commun.* **180** (2009) 120–156, [`0804.3755`].

[75] Particle Data Group collaboration, R. L. Workman et al., *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.

[76] J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, *Phys. Rev. Lett.* **100** (2008) 242001, [`0802.2470`].

[77] G. Luisoni, P. Nason, C. Oleari and F. Tramontano, *$HW^{\pm}$/HZ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO*, *JHEP* **10** (2013) 083, [`1306.2542`].

[78] F. Granata, J. M. Lindert, C. Oleari and S. Pozzorini, *NLO QCD+EW predictions for HV and HV +jet production including parton-shower effects*, *JHEP* **09** (2017) 012, [`1706.03522`].

[79] B. Hespel, F. Maltoni and E. Vryonidou, *Higgs and Z boson associated production via gluon fusion in the SM and the 2HDM*, *JHEP* **06** (2015) 065, [`1503.01656`].

[80] D. Goncalves, F. Krauss, S. Kuttimalai and P. Maierhöfer, *Higgs-Strahlung: Merging the NLO Drell-Yan and Loop-Induced 0+1 jet Multiplicities*, *Phys. Rev. D* **92** (2015) 073006, [`1509.01597`].

[81] S. Alioli, A. Broggio, S. Kallweit, M. A. Lim and L. Rottoli, *Higgsstrahlung at NNLL′+NNLO matched to parton showers in GENEVA*, *Phys. Rev. D* **100** (2019) 096016, [`1909.02026`].

[82] G. Ferrera, G. Somogyi and F. Tramontano, *Associated production of a Higgs boson decaying into bottom quarks at the LHC in full NNLO QCD*, *Phys. Lett. B* **780** (2018) 346–351, [`1705.10304`].

[83] C. Anastasiou, F. Herzog and A. Lazopoulos, *The fully differential decay rate of a Higgs boson to bottom-quarks at NNLO in QCD*, *JHEP* **03** (2012) 035, [`1110.2368`].

[84] V. Del Duca, C. Duhr, G. Somogyi, F. Tramontano and Z. Trocsanyi, *Higgs boson decay into b-quarks at NNLO accuracy*, *JHEP* **04** (2015) 036, [`1501.07226`].

[85] S. Alioli, A. Broggio, A. Gavardi, S. Kallweit, M. A. Lim, R. Nagar, D. Napoletano and L. Rottoli, *Resummed predictions for hadronic Higgs boson decays*, *JHEP* **04** (2021) 254, [`2009.13533`].

[86] Y. Hu, C. Sun, X.-M. Shen and J. Gao, *Hadronic decays of Higgs boson at NNLO matched with parton shower*, *JHEP* **08** (2021) 122, [`2101.08916`].

[87] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen and P. Z. Skands, *An Introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159–177, [`1410.3012`].

[88] 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`].

[89] O. Brein, R. Harlander, M. Wiesemann and T. Zirke, *Top-Quark Mediated Effects in Hadronic Higgs-Strahlung*, *Eur. Phys. J. C* **72** (2012) 1868, [`1111.0761`].

[90] J. M. Campbell, R. Ellis, R. Frederix, P. Nason, C. Oleari and C. Williams, *NLO Higgs Boson Production Plus One and Two Jets Using the POWHEG BOX, MadGraph4 and MCFM*, *JHEP* **07** (2012) 092, [`1202.5475`].

[91] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni, T. Plehn, D. L. Rainwater and T. Stelzer, *MadGraph/MadEvent v4: The New Web Generation*, *JHEP* **09** (2007) 028, [`0706.2334`].

[92] F. Granata, *Electroweak and strong next-to-leading-order corrections to HV and HVj production at hadron colliders*. PhD thesis, Milan Bicocca U., 2017.

[93] T. Gehrmann and E. Remiddi, *Numerical evaluation of harmonic polylogarithms*, *Comput. Phys. Commun.* **141** (2001) 296–312, [`hep-ph/0107173`].

[94] P. Nason and C. Oleari, *Generation cuts and Born suppression in POWHEG*, `1303.3922`.

[95] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001, [`hep-ph/9707323`].

[96] S. Bentvelsen and I. Meyer, *The Cambridge jet algorithm: Features and applications*, *Eur. Phys. J. C* **4** (1998) 623–629, [`hep-ph/9803322`].

[97] J. M. Campbell, R. K. Ellis, P. Nason and E. Re, *Top-Pair Production and Decay at NLO Matched with Parton Showers*, *JHEP* **04** (2015) 114, [`1412.1828`].

[98] R. Boughezal, J. M. Campbell, R. K. Ellis, C. Focke, W. Giele, X. Liu, F. Petriello and C. Williams, *Color singlet production at NNLO in MCFM*, *Eur. Phys. J. C* **77** (2017) 7, [`1605.08011`].

[99] A. Banfi, G. P. Salam and G. Zanderighi, *Infrared safe definition of jet flavor*, *Eur. Phys. J. C* **47** (2006) 113–124, [`hep-ph/0601139`].

[100] ATLAS collaboration, G. Aad et al., *Measurement of the associated production of a Higgs boson decaying into b-quarks with a vector boson at high transverse momentum in $pp$ collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Lett. B* **816** (2021) 136204, [`2008.02508`].

[101] NNPDF collaboration, R. D. Ball et al., *Parton distributions from high-precision collider data*, *Eur. Phys. J. C* **77** (2017) 663, [`1706.00428`].

[102] P. Skands, S. Carrazza and J. Rojo, *Tuning PYTHIA 8.1: the Monash 2013 Tune*, *Eur. Phys. J. C* **74** (2014) 3024, [`1404.5630`].

[103] LHC Higgs Cross Section Working Group collaboration, D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, `1610.07922`.

[104] M. Cacciari, G. P. Salam and G. Soyez, *The anti-$k_t$ jet clustering algorithm*, *JHEP* **04** (2008) 063, [`0802.1189`].

[105] S. Mrenna and P. Skands, *Automated Parton-Shower Variations in Pythia 8*, *Phys. Rev. D* **94** (2016) 074005, [`1605.08352`].

[106] ATLAS collaboration, M. Aaboud et al., *Observation of $H \to b\bar{b}$ decays and $VH$ production with the ATLAS detector*, *Phys. Lett. B* **786** (2018) 59–86, [`1808.08238`].

[107] CMS collaboration, A. M. Sirunyan et al., *Observation of Higgs boson decay to bottom quarks*, *Phys. Rev. Lett.* **121** (2018) 121801, [`1808.08242`].

[108] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock and B. R. Webber, *New clustering algorithm for multi - jet cross-sections in e+ e- annihilation*, *Phys. Lett. B* **269** (1991) 432–438.

[109] 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) 187–224.

[110] A. Banfi, G. P. Salam and G. Zanderighi, *Accurate QCD predictions for heavy-quark jets at the Tevatron and LHC*, *JHEP* **07** (2007) 026, [`0704.2999`].

[111] Super-Kamiokande collaboration, Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562–1567, [`hep-ex/9807003`].

[112] SNO collaboration, Q. R. Ahmad et al., *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **89** (2002) 011301, [`nucl-ex/0204008`].

[113] Planck collaboration, N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6, [`1807.06209`].

[114] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652–657.

[115] T. Appelquist and J. Carazzone, *Infrared Singularities and Massive Fields*, *Phys. Rev. D* **11** (1975) 2856.

[116] W. Buchmuller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, *Nucl. Phys. B* **268** (1986) 621–653.

[117] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085, [`1008.4884`].

[118] I. Brivio and M. Trott, *The Standard Model as an Effective Field Theory*, *Phys. Rept.* **793** (2019) 1–98, [`1706.08945`].

[119] K. Mimasu, V. Sanz and C. Williams, *Higher Order QCD predictions for Associated Higgs production with anomalous couplings to gauge bosons*, *JHEP* **08** (2016) 039, [`1512.02572`].

[120] C. Degrande, B. Fuks, K. Mawatari, K. Mimasu and V. Sanz, *Electroweak Higgs boson production in the standard model effective field theory beyond leading order in QCD*, *Eur. Phys. J. C* **77** (2017) 262, [`1609.04833`].

[121] R. Gauld, B. D. Pecjak and D. J. Scott, *One-loop corrections to $h \to b\bar{b}$ and $h \to \tau\bar{\tau}$ decays in the Standard Model Dimension-6 EFT: four-fermion operators and the large-$m_t$ limit*, *JHEP* **05** (2016) 080, [`1512.02508`].

[122] R. Gauld, B. D. Pecjak and D. J. Scott, *QCD radiative corrections for $h \to b\bar{b}$ in the Standard Model Dimension-6 EFT*, *Phys. Rev. D* **94** (2016) 074045, [`1607.06354`].

[123] 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`].

[124] J. M. Cullen and B. D. Pecjak, *Higgs decay to fermion pairs at NLO in SMEFT*, *JHEP* **11** (2020) 079, [`2007.15238`].

[125] W. Bizoń, F. Caola, K. Melnikov and R. Röntsch, *Anomalous couplings in associated VH production with Higgs boson decay to massive b quarks at NNLO in QCD*, *Phys. Rev. D* **105** (2022) 014023, [`2106.06328`].

[126] G. F. Giudice, C. Grojean, A. Pomarol and R. Rattazzi, *The Strongly-Interacting Light Higgs*, *JHEP* **06** (2007) 045, [`hep-ph/0703164`].

[127] J. Elias-Miro, J. R. Espinosa, E. Masso and A. Pomarol, *Higgs windows to new physics through d=6 operators: constraints and one-loop anomalous dimensions*, *JHEP* **11** (2013) 066, [`1308.1879`].

[128] G. D'Ambrosio, G. F. Giudice, G. Isidori and A. Strumia, *Minimal flavor violation: An Effective field theory approach*, *Nucl. Phys. B* **645** (2002) 155–187, [`hep-ph/0207036`].

[129] F. Caola, K. Melnikov and R. Röntsch, *Analytic results for decays of color singlets to gg and $q\bar{q}$ final states at NNLO QCD with the nested soft-collinear subtraction scheme*, *Eur. Phys. J. C* **79** (2019) 1013, [`1907.05398`].

[130] S. G. Gorishnii, A. L. Kataev and S. A. Larin, *The Width of Higgs Boson Decay Into Hadrons: Three Loop Corrections of Strong Interactions*, *Sov. J. Nucl. Phys.* **40** (1984) 329–334.

[131] S. G. Gorishnii, A. L. Kataev, S. A. Larin and L. R. Surguladze, *Corrected Three Loop QCD Correction to the Correlator of the Quark Scalar Currents and $\gamma$ (Tot) ($H^0 \to$ Hadrons)*, *Mod. Phys. Lett. A* **5** (1990) 2703–2712.

[132] S. G. Gorishnii, A. L. Kataev, S. A. Larin and L. R. Surguladze, *Scheme dependence of the next to next-to-leading QCD corrections to Gamma(tot) (H0 —> hadrons) and the spurious QCD infrared fixed point*, *Phys. Rev. D* **43** (1991) 1633–1640.

[133] K. G. Chetyrkin, *Correlator of the quark scalar currents and Gamma(tot) (H —> hadrons) at O (alpha-s**3) in pQCD*, *Phys. Lett. B* **390** (1997) 309–317, [`hep-ph/9608318`].

[134] P. A. Baikov, K. G. Chetyrkin and J. H. Kuhn, *Scalar correlator at O(alpha(s)**4), Higgs decay into b-quarks and bounds on the light quark masses*, *Phys. Rev. Lett.* **96** (2006) 012003, [`hep-ph/0511063`].

[135] F. Herzog, B. Ruijl, T. Ueda, J. A. M. Vermaseren and A. Vogt, *On Higgs decays to hadrons and the R-ratio at $N^4LO$*, *JHEP* **08** (2017) 113, [`1707.01044`].

[136] A. Hayreter and G. Valencia, *Constraints on anomalous color dipole operators from Higgs boson production at the LHC*, *Phys. Rev. D* **88** (2013) 034033, [`1304.6976`].

[137] J. Bramante, A. Delgado, L. Lehman and A. Martin, *Boosted Higgses from chromomagnetic b's: $b\bar{b}h$ at high luminosity*, *Phys. Rev. D* **93** (2016) 053001, [`1410.3484`].

[138] U. Haisch and G. Koole, *Beautiful and charming chromodipole moments*, *JHEP* **09** (2021) 133, [`2106.01289`].

[139] 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`].

[140] Particle Data Group collaboration, P. Zyla et al., *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01.

[141] ATLAS collaboration, G. Aad et al., *Measurements of the production cross-section for a Z boson in association with b-jets in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **07** (2020) 044, [`2003.11960`].

[142] ATLAS collaboration, G. Aad et al., *Measurements of $WH$ and $ZH$ production in the $H \to b\bar{b}$ decay channel in $pp$ collisions at 13 TeV with the ATLAS detector*, *Eur. Phys. J. C* **81** (2021) 178, [`2007.02873`].

[143] M. Grazzini, S. Kallweit, J. M. Lindert, S. Pozzorini and M. Wiesemann, *NNLO QCD + NLO EW with Matrix+OpenLoops: precise predictions for vector-boson pair production*, *JHEP* **02** (2020) 087, [`1912.00068`].

[144] J. Baglio, L. D. Ninh and M. M. Weber, *Massive gauge boson pair production at the lhc: A next-to-leading order story*, *Phys. Rev. D* **88** (Dec, 2013) 113005.

[145] J. Ohnemus, *An Order $\alpha^{-}s$ calculation of hadronic $W^{\pm}Z$ production*, *Phys. Rev. D* **44** (1991) 3477–3489.

[146] J. Ohnemus, *Hadronic ZZ, $W^{-}W^{+}$, and $W^{\pm}Z$ production with QCD corrections and leptonic decays*, *Phys. Rev. D* **50** (1994) 1931–1945, [`hep-ph/9403331`].

[147] J. M. Campbell and R. K. Ellis, *An Update on vector boson pair production at hadron colliders*, *Phys. Rev. D* **60** (1999) 113006, [`hep-ph/9905386`].

[148] L. J. Dixon, Z. Kunszt and A. Signer, *Vector boson pair production in hadronic collisions at order $\alpha_s$ : Lepton correlations and anomalous couplings*, *Phys. Rev. D* **60** (1999) 114037, [`hep-ph/9907305`].

[149] J. M. Campbell, R. K. Ellis and C. Williams, *Vector boson pair production at the LHC*, *JHEP* **07** (2011) 018, [`1105.0020`].

[150] A. Denner and G. Pelliccioli, *NLO QCD predictions for doubly-polarized WZ production at the LHC*, *Phys. Lett. B* **814** (2021) 136107, [`2010.07149`].

[151] F. Campanario, C. Englert, S. Kallweit, M. Spannowsky and D. Zeppenfeld, *NLO QCD corrections to WZ+jet production with leptonic decays*, *JHEP* **07** (2010) 076, [`1006.0390`].

[152] M. Grazzini, S. Kallweit, D. Rathlev and M. Wiesemann, *$W^{\pm}Z$ production at hadron colliders in NNLO QCD*, *Phys. Lett. B* **761** (2016) 179–183, [`1604.08576`].

[153] M. Grazzini, S. Kallweit, D. Rathlev and M. Wiesemann, *$W^{\pm}Z$ production at the LHC: fiducial cross sections and distributions in NNLO QCD*, *JHEP* **05** (2017) 139, [`1703.09065`].

[154] M. Grazzini, S. Kallweit and M. Wiesemann, *Fully differential NNLO computations with MATRIX*, *Eur. Phys. J. C* **78** (2018) 537, [`1711.06631`].

[155] J. M. Campbell, R. K. Ellis and S. Seth, *Non-local slicing approaches for NNLO QCD in MCFM*, *JHEP* **06** (2022) 002, [`2202.07738`].

[156] A. Bierweiler, T. Kasprzik and J. H. Kühn, *Vector-boson pair production at the LHC to $\mathcal{O}(\alpha^3)$ accuracy*, *JHEP* **1312** (2013) 071, [`1305.5402`].

[157] J. Baglio, L. D. Ninh and M. M. Weber, *Massive gauge boson pair production at the LHC: a next-to-leading order story*, *Phys. Rev. D* **88** (2013) 113005, [`1307.4331`].

[158] B. Biedermann, A. Denner and L. Hofer, *Next-to-leading-order electroweak corrections to the production of three charged leptons plus missing energy at the LHC*, *JHEP* **10** (2017) 043, [`1708.06938`].

[159] M. Chiesa, C. Oleari and E. Re, *NLO QCD+NLO EW corrections to diboson production matched to parton shower*, *Eur. Phys. J. C* **80** (2020) 849, [`2005.12146`].

[160] E. Bothmann, D. Napoletano, M. Schönherr, S. Schumann and S. L. Villani, *Higher-order EW corrections in ZZ and ZZj production at the LHC*, *JHEP* **06** (2022) 064, [`2111.13453`].

[161] ATLAS collaboration, M. Aaboud et al., *Measurement of $W^{\pm}Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **79** (2019) 535, [`1902.05759`].

[162] R. W. Brown, D. Sahdev and K. O. Mikaelian, *W+- Z0 and W+- gamma Pair Production in Neutrino e, p p, and anti-p p Collisions*, *Phys. Rev. D* **20** (1979) 1164.

[163] F. Cascioli, P. Maierhöfer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601, [`1111.5206`].

[164] F. Buccioni, S. Pozzorini and M. Zoller, *On-the-fly reduction of open loops*, *Eur. Phys. J. C* **78** (2018) 70, [`1710.11452`].

[165] F. Buccioni, J.-N. Lang, J. M. Lindert, P. Maierhöfer, S. Pozzorini, H. Zhang and M. F. Zoller, *OpenLoops 2*, *Eur. Phys. J. C* **79** (2019) 866, [`1907.13071`].

[166] T. Gehrmann, A. von Manteuffel and L. Tancredi, *The two-loop helicity amplitudes for* $q\overline{q}' \to V_1V_2 \to 4$ *leptons*, *JHEP* **09** (2015) 128, [`1503.04812`].

[167] A. Manohar, P. Nason, G. P. Salam and G. Zanderighi, *How bright is the proton? A precise determination of the photon parton distribution function*, *Phys. Rev. Lett.* **117** (2016) 242002, [`1607.04266`].

[168] A. V. Manohar, P. Nason, G. P. Salam and G. Zanderighi, *The Photon Content of the Proton*, *JHEP* **12** (2017) 046, [`1708.01256`].

[169] NNPDF collaboration, V. Bertone, S. Carrazza, N. P. Hartland and J. Rojo, *Illuminating the photon content of the proton within a global PDF analysis*, *SciPost Phys.* **5** (2018) 008, [`1712.07053`].

[170] A. Denner, S. Dittmaier, M. Roth and D. Wackeroth, *Predictions for all processes e+ e- —> 4 fermions + gamma*, *Nucl. Phys. B* **560** (1999) 33–65, [`hep-ph/9904472`].

[171] 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) 247–294, [`hep-ph/0505042`].

[172] ATLAS collaboration, M. Aaboud et al., *Measurement of the* $W^{\pm}Z$ *boson pair-production cross section in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS Detector*, *Phys. Lett.* **B762** (2016) 1–22, [`1606.04017`].

[173] C. Bierlich et al., *Robust Independent Validation of Experiment and Theory: Rivet version 3*, *SciPost Phys.* **8** (2020) 026, [`1912.05451`].

[174] C. M. Carloni Calame, G. Montagna, O. Nicrosini and A. Vicini, *Precision electroweak calculation of the production of a high transverse-momentum lepton pair at hadron colliders*, *JHEP* **10** (2007) 109, [`0710.1722`].

[175] S. Dittmaier and M. Huber, *Radiative corrections to the neutral-current Drell-Yan process in the Standard Model and its minimal supersymmetric extension*, *JHEP* **01** (2010) 060, [`0911.2329`].

[176] S. Kallweit, J. M. Lindert, P. Maierhöfer, S. Pozzorini and M. Schönherr, *NLO QCD+EW predictions for V + jets including off-shell vector-boson decays and multijet merging*, *JHEP* **04** (2016) 021, [`1511.08692`].