Automation of NLO QCD Corrections and the Application to N-Jet Processes at Lepton Colliders

Dissertation zur Erlangung des Doktorgrades an der Fakultät für Mathematik, Informatik und Naturwissenschaften
Fachbereich Physik
der Universität Hamburg

vorgelegt von
Vincent Rothe

Hamburg
2021

Gutachter der Dissertation:	Dr. Jürgen Reuter Prof. Dr. Sven-Olaf Moch
Zusammensetzung der Prüfungskommission:	Dr. Jürgen Reuter Prof. Dr. Sven-Olaf Moch Prof. Dr. Caren Hagner Dr. Markus Diehl Prof. Dr. Peter Schleper
Vorsitzende/r der Prüfungskommission:	Prof. Dr. Caren Hagner
Datum der Disputation:	08.07.2021
Vorsitzender Fach-Promotionsausschuss PHYSIK:	Prof. Dr. Wolfgang Hansen
Leiter des Fachbereichs PHYSIK:	Prof. Dr. Günter H. W. Sigl
Dekan der Fakultät MIN:	Prof. Dr. Heinrich Graener

Abstract

Despite strong indications for physics beyond the Standard Model (BSM), no definite direct detection of new physics has been made at the Large Hadron Collider (LHC) yet. Therefore, high-precision particle physics becomes ever more important in order to probe the Standard Model (SM) parameters for any deviation from theoretical predictions.

This requires increasing experimental precision and theoretical accuracy. On the experimental side – besides improving hardware, luminosities, etc. – lepton colliders such as the planned International Linear Collider (ILC) provide very clean colorless initial states within a triggerless detection mode and eliminate uncertainties from non-perturbative hadronic parton distribution functions.

Theoretical accuracy can be improved by computing higher-order corrections in the perturbation series beyond leading order (LO). Monte Carlo (MC) event generators automate these computations by the implementation of subtraction schemes to treat infrared divergences that occur at next-to-leading order (NLO) and beyond.

The multi-purpose MC event generator WHIZARD has a wide feature set specific to lepton colliders and is well-established in the community. In this thesis, we present our work on generalizing its NLO quantum chromodynamics (QCD) capabilities for arbitrary processes at fixed order using the Frixione-Kunszt-Signer (FKS) subtraction scheme. This work is based on a prior FKS-subtraction implementation which was limited to a narrow range of processes.

First, we give a thorough review of FKS subtraction, followed by a detailed description of its implementation in WHIZARD. We then validate our fixed-NLO (fNLO) QCD implementation by carrying out many self-consistency checks and by performing an intensive comparison of WHIZARD results with two other event generators at the level of total cross sections and differential distributions.

As far as we know, we compute the first result for the total cross section of 6-jet production at fNLO QCD at a lepton collider with $\sigma_{\mathrm{NLO}}^{e^+e^-\to jjjjjj} = 4.46(4)$ fb.

Zusammenfassung

Mit ausbleibenden Entdeckungen von Physik jenseits des Standard Modells am Large Hadron Collider (LHC) wird die Bestimmung von Parametern des Standard Modells mit steigender Präzision immer wichtiger, um Abweichungen von den Vorhersagen – und damit neue Physik – nachzuweisen.

Aus Sicht der Experimentalphysik bieten Leptonenbeschleuniger (z.B. der geplante International Linear Collider (ILC)) ideale Verhältnisse für Hochpräzisionsmessungen aufgrund des geringeren Untergrundes durch Quantenchromodynamik (QCD) und die nicht notwendigen hadronischen Partondichtefunktionen, die nicht störungstheoretisch berechnet werden können.

Höhere theoretische Genauigkeit bietet das Berechnen von Termen in der Störungsreihe jenseits der führenden Ordnung (LO) in den Kopplungskonstanten. Dessen Automatisierung in Monte Carlo (MC) Eventgeneratoren ermöglichen Subtraktionsschemata, die Infrarot-Divergenzen in nächst-höherer Ordnung (NLO) getrennt subtrahieren.

Der Eventgenerator WHIZARD ist etabliert für Studien an Leptonenbeschleunigern, da er hierfür eine breite Palette an spezialisierten Funktionen liefert. Eine erste Implementation des Frixione-Kunszt-Signer (FKS) Subtraktionsschemas in WHIZARD ermöglichte NLO QCD Berechnungen in fester Ordnung der Störungstheorie für wenige Prozesse, deren Erweiterung für beliebige Prozesse wir in dieser Arbeit präsentieren.

Wir fassen die Theorie zur FKS Subtraktion zusammen und detailieren dessen Implementation in WHIZARD. Diese validieren wir umfangreich, indem wir zunächst eine Reihe von Selbsttests ausfhren, um die technische Integrität unserer Implementation in WHIZARD sicherzustellen. Daraufhin führen wir für eine große Auswahl an Prozessen an Leptonen- und Hadronenbeschleunigern auf fester-NLO (fNLO) Genauigkeit in QCD einen ausführlichen und tiefgehenden Vergleich mit zwei weiteren Eventgeneratoren durch. Damit validieren wir die Fähigkeit WHIZARDs, für beliebige Prozesse vollautomatisch fNLO QCD Korrekturen zu totalen Wirkungsquerschnitten sowie differentiellen Verteilungen zu berechnen.

Soweit uns bekannt ist, berechnen wir das erste Resultat für den totalen Wirkungsquerschnitt von 6-Jetproduktion auf fNLO QCD Genauigkeit an einem Leptonenbeschleuniger mit $\sigma_{\mathrm{NLO}}^{e^+e^-\to jjjjjj} = 4.46(4)$ fb.

List of Publications

Publications and proceedings that I have contributed to and that are related to this thesis:

[1] P. Stienemeier, S. Braß, P. Bredt, et al. "WHIZARD 3.0: Status and News". *International Workshop on Future Linear Colliders*. 2021. arXiv: 2104.11141 [hep-ph].

[2] J. Reuter, S. Braß, P. Bredt, et al. "Status of the WHIZARD generator for linear colliders". *International Workshop on Future Linear Colliders*. 2020. arXiv: 2002.06122 [hep-ph].

[3] S. Braß, W. Kilian, T. Ohl, et al. *Precision Monte Carlo simulations with WHIZARD*. CERN Yellow Reports: Monographs, **3** (2020). Ed. by A. Blondel, J. Gluza, S. Jadach, P. Janot, and T. Riemann, pp. 205–210. DOI: 10.23731/CYRM-2020-003.205.

[4] W. Kilian, S. Braß, T. Ohl, et al. "New Developments in WHIZARD Version 2.6". *International Workshop on Future Linear Collider*. 2018. arXiv: 1801.08034 [hep-ph].

Publications still in preparation that I have contributed to and that are related to this thesis:

[1] S. Braß, P. Bredt, W. Kilian, et al. "Complete Automation of NLO SM Processes in WHIZARD". in preparation.

Publications and proceedings that I have contributed to during my PhD studies that are not directly related to this thesis:

[1] A. Ballestrero et al. *Precise predictions for same-sign W-boson scattering at the LHC*. Eur. Phys. J. C, **78**.8 (2018), p. 671. DOI: 10.1140/epjc/s10052-018-6136-y. arXiv: 1803.07943 [hep-ph].

[2] C. F. Anders et al. *Vector boson scattering: Recent experimental and theory developments*. Rev. Phys., **3** (2018), pp. 44–63. DOI: 10.1016/j.revip.2018.11.001. arXiv: 1801.04203 [hep-ph].

List of Acronyms

BLHA Binoth Les Houches Accord
BSM beyond the Standard Model

CL confidence level
CLIC Compact Linear Collider
CM center-of-mass
CS Catani-Seymour

DESY Deutsches Elektronen-Synchrotron
DGLAP Dokshitzer-Gribov-Lipatov-Altarelli-Parisi

EW electroweak

FKS Frixione-Kunszt-Signer
fLO fixed-LO
fNLO fixed-NLO
FO fixed-order
FSR final-state radiation

HEP high energy physics

ILC International Linear Collider
IR infrared
ISR initial-state radiation

KLN Kinoshita-Lee-Nauenberg

LEP Large Electron-Positron Collider
LHC Large Hadron Collider
LO leading order

MC Monte Carlo
MCI multi-channel integrator
MPI message passing interface

NLO next-to-leading order
NNLO next-to-next-to-leading order

OLP one-loop provider

PDF parton distribution function
PDG Particle Data Group

QCD quantum chromodynamics
QED quantum electrodynamics
QFT quantum field theory

RNG random number generator

SM Standard Model

UV ultraviolet

List of Figures

List of Tables

Contents

1. Introduction

The last decade of particle physics saw a great success with the discovery of a particle that has long been predicted by the Standard Model (SM) of particle physics – the Higgs boson. Its existence follows from the mechanism that gives elementary particles their masses, namely electroweak (EW) symmetry breaking [1–6]. In 2012, a Higgs-like state with a mass of 125 GeV was discovered by the ATLAS [7] and CMS [8] experiments at the Large Hadron Collider (LHC) [9], concluding a decades-long search.

In addition to the Higgs boson, many new hadrons have been discovered in the data that was collected in the first two runs of the LHC at a center-of-mass (CM) energy of 7 to 8 TeV (2009–2013) and 13 TeV (2015–2018). For example, the LHCb experiment reported observed resonances consistent with pentaquark states [10]. However, all of these observations are in agreement with the SM and while the LHC continues to improve exclusion limits for searches beyond the Standard Model (BSM), expectations for significant discoveries of new physics have not yet been met. This is contrary to observational evidence for physics that cannot be described by the SM alone, like the most popular example of dark matter and dark energy. We know that ordinary baryonic matter and energy only contribute a small fraction to the total energy density of the universe. This knowledge comes, for example, from the observations of the rotation curves of galaxies [11, 12], estimates of masses for galaxy clusters in relation to their visible matter [13], gravitational lensing [14, 15], and the precise measurement of the cosmic microwave background [16].

In the case of dark matter, a multitude of different theoretical explanations exist, many of which involve one or more new particles (see [17] for a detailed review). The general consensus is that they must only interact weakly (if at all) with SM physics. In order to explain the observations listed above, however, all theories must allow for gravitational interaction.

Another promising hint for new physics is given by the recent high precision measurement by the Muon $g-2$ experiment at Fermilab of the positive muon anomalous magnetic moment which is found to deviate from the SM prediction by a significance of 4.2σ [18]. These results could indicate SM extensions with new couplings to leptons. It remains to be seen if future runs of the experiment increase the significance.

With the lack of direct evidence for BSM physics at the LHC at increasing energies and luminosities, more effort in the high energy physics (HEP) community is being

directed towards high-precision studies. These studies try to significantly reduce uncertainties in crucial SM parameters, such as couplings, particle masses, mixing angles, and CP violation parameters. In return, more precise results can then indirectly lead to new discoveries as they increase the accuracy of the models used in the analyses. Therefore, deviations between SM predictions and data that may have been buried by larger uncertainties before might become significant enough to be regarded as a discovery.[1]

Increasing precision is a two-sided story, however. On the one hand, improved or new experimental methods, increased sensitivity and capabilities of detector hardware, as well as advances in computing power, have brought and continue to bring down systematic uncertainties significantly. Statistical errors can be reduced by increasing the luminosity of a particle collider, e.g. as planned for both the High-Luminosity and High-Energy LHC [19, 20]. Furthermore, future lepton colliders – such as the International Linear Collider (ILC) [21], the Compact Linear Collider (CLIC) [22], or the lepton-collider option of the Future Circular Collider (FCC) [23, 24] – provide an excellent environment for high-precision studies. In comparison to hadron colliders, their initial state is clean of any quantum chromodynamics (QCD) radiation which dramatically reduces the overall background and thus uncertainties. Additionally, while at hadron colliders the colliding partons carry a random fraction of the original beam energy, the energy of the hard interaction can be much more fine-tuned at lepton colliders. Initial-state quantum electrodynamics (QED) effects such as beamstrahlung and initial-state radiation (ISR) of photons still have to be accounted for, however.

On the other hand, in order to match the smaller experimental uncertainties, theoretical calculations must reach a higher precision to provide an improved comparison between predictions for both signal and background processes and the experimental observations. In quantum field theory (QFT), the matrix element for the hard interaction – i.e. the underlying scattering process at the highest energy scale – is approximatively described by a perturbative series in the coupling constant of the corresponding theory. We call this an fixed-order (FO) expansion. By computing terms of increasing powers in the coupling constant, the total cross section σ of the process in question receives higher-order corrections. For example, in QCD we have [25]

$$\sigma = \sigma_0 + \alpha_S \sigma_1 + \alpha_S^2 \sigma_2 + \alpha_S^3 \sigma_3 + \cdots , \tag{1.0.1}$$

with the strong coupling constant α_S and the leading order (LO) (or more precisely

[1] The generally excepted threshold for a discovery is a significance of 5σ.

Born) cross section σ_0.[2] Each subsequent term σ_k with an explicit order in the coupling constant contributes as the (next-to)^k-leading order (N^{k}LO) correction to the LO cross section and thus improves the precision of the total cross section. For each additional power of α_S, Feynman diagrams have to be added that correspond to the real radiation of an additional particle that contributes to the number of particles in the final state, as well as diagrams that describe the virtual loop corrections due to the radiation and reabsorption of an additional off-shell particle.

During the calculation of higher-order corrections, divergences start to appear at different scales of the energy spectrum. Beginning with the evaluation of virtual loop corrections, for every additional power in the coupling constant we find Feynman graphs with closed loops that give rise to integrals over all values of unconstrained loop momenta. At high energy scales these integrals induce so-called ultraviolet (UV) divergences. Because of the fact that we require all QFTs that describe the SM to be renormalizable, the corresponding UV singularities can be treated by an absorption into a finite number of parameters and fields entering the Lagrangian of the theory, e.g. the coupling constants. This introduces a dependence of the coupling constants on an arbitrary energy scale μ_R – the so-called *renormalization scale*. When all orders in the perturbation series are included, physical observables do not depend on μ_R, as it is canceled out in the calculation. For computations up to a fixed order in the coupling constant, observables remain dependent on μ_R. Variations around a chosen renormalization scale are commonly used to estimate the theoretical uncertainties which arise by not calculating the missing higher-order contributions.

With the UV divergences taken care of, we still have to deal with divergences that arise at the very low scales of the energy spectrum, as well as divergences due to the momenta of a radiated particle and its emitter becoming collinear. It is common to combine both the soft (i.e. low-energy) and collinear divergences under the term infrared (IR) divergences.[3] These divergences are present in both the calculations of loop and real-emission diagrams and can best be visualized in QCD for an emission of a gluon from a quark by looking at the denominator of the quark propagator. It can be written in terms of the four-momenta q and g of the quark and gluon, respectively, after the splitting as [27]

$$\frac{1}{(q+g)^2 - m_q^2} = \frac{1}{2q^0 g^0 \left(1 - \beta_q \cos\Theta_{qg}\right)}, \quad \beta_q \equiv \frac{|\mathbf{q}|}{q^0} = \sqrt{1 - \frac{m_q^2}{\left(q^0\right)^2}}, \tag{1.0.2}$$

[2] There are subtle to significant differences between the LO cross section and the Born contribution in a higher-order calculation. The latter will generally be computed for particle widths at the corresponding order, as well as for different scales compared to the LO computation [26].

[3] Sometimes IR may only refer to the soft divergences, with explicit mentions of the collinear divergences. We include both in the term IR.

with the mass m_q and three-vector $\mathbf{q}$ of the quark, and the angle Θ_{qg} between the quark and the gluon. For the radiation of a very soft gluon with an energy $g^0 \to 0$, the propagator becomes singular, showcasing the soft divergence.[4] Regardless of the gluon (or the quark) becoming soft, we also encounter a divergent behavior if Θ_{qg} becomes small, i.e. if the gluon becomes collinear to the quark. This is only true for the radiation off massless particles, however, as otherwise the mass in the factor β_q regularizes the collinear singularity.

In order to regularize the integrals over the IR singularities, it is common to use the technique of *dimensional regularization*, where the $d = 4$ dimensions of Minkowski space (one energy and three spatial dimensions) are analytically continued to $d = 4-2\epsilon$ dimensions, with $\epsilon < 0$. Evaluation of the phase-space integrals then results in expressions for the real-emission and virtual amplitudes dependent on $1/\epsilon$ and $1/\epsilon^2$ poles that cancel each other when both results are added to give the total higher-order correction. Subsequently, the limit $\epsilon \to 0$ can then be taken, resulting in a finite result in $d = 4$ dimensions. This observation of the cancellation of IR divergences has led to the Kinoshita-Lee-Nauenberg (KLN) theorem [28, 29]. It is a generalization of the much earlier Bloch-Nordsieck theorem [30] for QED and states that in a theory with massless fields (or in its massless limit) any IR divergences cancel for each order of the perturbation series, given that the calculation is done for sufficiently inclusive and IR-save observables, i.e. without sensitivity to momenta that are highly soft or collinear.

Theoretical predictions are crucial for the guidance in planning collider experiments, as well as for signal and background predictions in analyses of data obtained from these experiments. In principle, these predictions are necessary for any conceivable process that can be produced in the experiments. However, already at LO for increasing final-state multiplicities, this quickly becomes a task that is almost impossible to complete by analytical computations alone, as increasing numbers of complicated integrals in high-dimensional phase spaces have to be performed. Particle physics therefore has a long and successful history of employing a multitude of different computer programs in order to solve this task.

Monte Carlo (MC) event generators are especially suited for the evaluation of phase-space integrals of high dimensionality and event simulation (see [31] for a review). Many of them allow for a high level of automation and flexibility in calculating total cross sections and simulating events for almost arbitrary processes in the SM and beyond. All of these programs are at minimum capable to provide these calculations and simulations

[4] As already noted in [26], although this is commonly used as an example for a soft divergence, it is not sufficient for the proof of its existence as the phase-space volume could cancel the energy g^0 of the gluon. In section 2.2, when discussing soft counterterms, we will encounter a much more definite way of showing soft divergences.

at the partonic level of the hard process within the bounds of perturbation theory at fixed-LO (fLO). To accurately describe the full event detected in an experiment, however, further physical phenomena have to be taken into account, such as the successive radiation of soft and collinear partons by those stemming from the hard process, down to the energy scale at which hadrons start to form. This is called a *parton shower*, with the process of confining the QCD partons into hadrons being called *hadronization*. While many *multi-purpose* MC event generators – such as MG5_aMC@NLO [32, 33], Sherpa [34], or WHIZARD [35] – have their own implementations of these aspects of event simulation, MC programs with a stronger focus on parton showers, as well as hadronization and subsequent decay exist. Prominent examples are Pythia [36, 37] and Herwig [38, 39], which can take the parton level event samples produced by other MC event generators as input and process them accordingly.

At orders beyond LO, the difficulty lies in the fact that MC integration cannot be done in non-integer dimensions, meaning that we cannot rely on the dimensional regularization method used in the analytic computations. Therefore, at next-to-leading order (NLO), methods have been developed based on the KLN theorem to regularize the real and virtual contributions separately in order for them to be integrated in $d = 4$ dimensions. The prevailing methods used today are so-called *subtraction schemes*, most notably the Catani-Seymour (CS) [40] and Frixione-Kunszt-Signer (FKS) [41, 42] subtraction schemes. Additionally, the development of methods to compute one-loop integrals based on the reduction of tensor integrals to scalar master integrals [43–47] led to the advent of efficient one-loop providers (OLPs) like the matrix-element generators GoSam [48], MadLoop [49], NJet [50, 51], OpenLoops [52, 53], and RECOLA [54–57]. Together with these tools and the subtraction schemes, many MC event generators nowadays are able to provide semi- and fully-automated NLO computations. These programs include `HELAC` [58], MG5_aMC@NLO, `POWHEG BOX` [59], Sherpa, VBFNLO [60], and WHIZARD.

Studies of NLO corrections and their computation and implementation in MC simulations are far from being exhausted and many processes remain to be calculated at NLO, especially those with high multiplicities in their final state. Furthermore, with future lepton colliders on the horizon, NLO EW and mixed corrections will become increasingly important. These have so far been of lesser importance due to the dominance of QCD contributions at higher orders at the LHC.

Beyond NLO, fully automatic calculations remain unsolved. At next-to-next-to-leading order (NNLO), there are promising recent developments within e.g. `EERAD3` [61], Matrix [62], MCFM [63] and NNLOjet [64]. So far, these are restricted to a small class of processes. At N^{3}LO, solutions are even sparser. A notable example is the cross section for Higgs boson production via gluon fusion at the LHC [65].

The overall structure of this thesis is divided into two parts. Part I starts with

chapter 2, where we begin by giving a short general introduction to the ideas behind common subtraction schemes. This is followed by a thorough review of the FKS subtraction scheme. We summarize the various different aspects of the scheme introduced in many different sources [41, 42, 59, 66–69] into a single consistent overview, needed to understand the full fixed-NLO (fNLO) implementation in WHIZARD for QCD. In order to accomplish this, we first establish a generalized notation, before delving into the details of phase-space partitioning and the construction of real and virtual subtraction terms. We want to be flexible and to perform NLO QCD simulations for both lepton and hadron colliders. Therefore, we do not only cover the subtraction of final-state divergences, but also detail the construction of initial-state subtraction terms and the treatment of additional collinear divergences introduced by the presence of parton distribution functions (PDFs).

In general, the FKS subtraction scheme is not limited to QCD corrections, but can also be implemented for EW and mixed corrections as well. To correctly describe EW effects, however, the subtraction terms have to be adapted accordingly [70]. Effects from photon PDFs at hadron colliders and electron PDFs at lepton colliders (see e.g. [71, 72]) have to be taken into consideration as well. While we implemented these necessary parts in WHIZARD, they are still in development and are undergoing validation. Consequently, we focus only on QCD corrections and refrain from presenting the theory for EW FKS subtraction and its implementation in WHIZARD in this work.

In chapter 3, after a short general overview of WHIZARD, we give a detailed description of the aspects of WHIZARD that are relevant in the context of FKS subtraction. Building upon the work presented in two earlier PhD theses [26, 73] that introduced a first and basic implementation of FKS subtraction in WHIZARD – limited to lepton-collider processes that only allow for final-state radiation (FSR) of an additional gluon – we outline the various changes that were necessary for a fully generalized support of arbitrary fNLO QCD processes. These changes include several improvements e.g. in the internal structure of FKS subtraction terms in WHIZARD, the phase-space setup and combinatorics required for FKS subtraction, the handling of symmetry factors, as well as a general implementation for the correct application of structure functions such as PDFs to FKS subtraction terms that can be easily adapted for the EW case. Additionally, the handling of differential distributions at fNLO within the WHIZARD framework had to be generalized. A detailed description of fNLO events and their representation in WHIZARD concludes part I.

In part II, we give an extensive validation of the fNLO QCD capabilities of WHIZARD in chapter 4. We first report on self-consistency checks to ensure the program is working as intended on a technical level in section 4.1. In section 4.2, we then perform an extensive comparison study on fNLO QCD total cross sections computed by WHIZARD for a vast list of both lepton and hadron-collider processes where we

use MG5_aMC@NLO as a reference.

Given the success of our exhaustive comparison of the implementation, we apply the fNLO QCD capabilities of WHIZARD to multi-jet production at lepton colliders. These processes have always been of great interest, partly due to the fact that jets can be defined in an IR-safe way [74], allowing to compute jet-based observables for each order of α_S in the perturbation series (see eq. (1.0.1)). Since lepton collisions are free from initial-state QCD effects, these processes provide an excellent opportunity to conduct high-precision studies on e.g. hadronization models or SM parameters like the value of α_S. Especially on the latter, many studies have been performed in the past in the context of the Large Electron-Positron Collider (LEP), often with the help of *event-shape* observables, such as thrust [75, 76] and the Parisi C- and D-parameters [77–79]. These event shapes are sensitive to the radiation of additional partons and hence also to increasing orders of α_S^{n-2} in n-jet processes. For 2- and 3-jet final states analyses exist with up to $\mathrm{N^2LO}$ accuracy (see [80] for an overview), with NLO accuracy for 4-jet [81–83]. For 5-jet final states NLO computations have been done for LEP (e.g. [84]).

At a future lepton collider, like ILC or CLIC, multi-jet production will provide even more data at higher energies and luminosities for these high-precision studies. Moreover, n-jet processes are important as QCD corrections to the process $e^+e^- \to f\bar{f}$, as well as backgrounds in multi-boson production, both also relevant in BSM studies. Therefore, it is of paramount importance for contemporary MC event generators that will be used for future studies to be thoroughly checked against each other. Any deviations between them have to be traced down, and properly understood and documented if due to differences at the level of their respective physics implementation. We thus give an in-depth comparison study of differential distributions between MG5_aMC@NLO, Sherpa, and WHIZARD at fNLO QCD in section 4.3 for jet observables and event shapes for multiplicities of up to 5-jets in 1 TeV e^+e^--collisions. In order to provide an example for the capabilities of our FKS implementation for ISR, we also give a short comparison between differential distributions generated by WHIZARD and the other two event generators for the process $pp \to Zj$ at a CM energy of 13 TeV.

In chapter 5, we showcase the potential of studying multi-jet processes with multiplicities beyond 5-jet at fNLO QCD with WHIZARD, enabled by our FKS implementation and the excellent capability of WHIZARD to handle phase spaces of high-multiplicity final states. As far as we know, we calculate the first full fNLO QCD result of the 6-jet total cross section at a lepton collider.

We conclude the thesis in chapter 6 with a summary of the results that have been achieved and end with a discussion on future extensions of our FKS implementation for mixed-EW and NNLO corrections.

Part I.

NLO Computations in a Multi-Purpose Monte Carlo Event-Generator Framework

2. Frixione-Kunszt-Signer Subtraction Scheme at NLO QCD

The total cross section at NLO for an arbitrary hard $2 \to n$ scattering process, where we neglect initial-state effects like PDFs, can be written as

$$\sigma_{\text{NLO}} = \int d\Phi_n \mathcal{B}(\Phi_n) + \int d\Phi_{n+1} \mathcal{R}(\Phi_{n+1}) + \int d\Phi_n \mathcal{V}_b(\Phi_n)\,, \tag{2.0.1}$$

where we denote the squared amplitude of the Born matrix element by $\mathcal{B}$ and the squared real amplitude by $\mathcal{R}$, dependent on the n-body and $(n+1)$-body phase space Φ_n and Φ_{n+1}, respectively. The virtual contribution $\mathcal{V}_b$ is made up by the interference of the Born and one-loop matrix elements. We call it the *bare* virtual contribution due to the fact that while we assume UV divergences to be taken care of by the means of renormalization, IR divergences still remain [67].

In chapter 1, we pointed out that the integration in eq. (2.0.1) over the n-body and $(n+1)$-body phase space for the virtual and real contribution, respectively, is IR-divergent. Since MC integration can only be performed in integer dimensions, the method of dimensional regularization with $d = 4 - 2\epsilon$ dimensions cannot be used in a numerical computation. However, we can apply suitable subtraction (or counter-)terms $\mathcal{C}^{(\alpha)}(\Phi_{n+1})$ for each singular configuration α of Φ_{n+1} that cancel the poles individually for the real contribution such that the limit

$$\lim_{\epsilon \to 0} \int d\Phi_{n+1} \left[\mathcal{R}(\Phi_{n+1}) - \sum_{\alpha} \mathcal{C}^{(\alpha)}(\Phi_{n+1}) \right] \tag{2.0.2}$$

is finite. According to the KLN theorem (see chapter 1), we can assume that the singular structure of the ϵ poles is identical for the real and virtual contributions, except for opposing signs. Therefore, the limit

$$\lim_{\epsilon \to 0} \int d\Phi_n \left[\mathcal{V}_b(\Phi_n) + \sum_{\alpha} \int d\Phi_{\text{rad}}^{(\alpha)} \mathcal{C}^{(\alpha)}(\Phi_{n+1}) \right] \tag{2.0.3}$$

has to be finite as well when the integration over the phase space of the radiated particle is carried out in $d = 4-2\epsilon$ for $\mathcal{C}^{(\alpha)}$. Indeed, we can formulate this in a universal way, meaning that the functional form of the counterterms $\mathcal{C}^{(\alpha)}$ in eq. (2.0.2) and their integration in eq. (2.0.3) have to be calculated analytically only once, after which they can be applied to arbitrary processes.

The integral over the radiation phase space $\Phi_{\text{rad}}^{(\alpha)}$ in eq. (2.0.3) implies that we have a mapping between the n-body and $(n+1)$-body phase space that factorizes Φ_{n+1} in each α. However, this mapping and the construction of the counterterms is dependent on the representation of the singular configuration α. Therefore, the $\mathcal{C}^{(\alpha)}$ are not unique, resulting in different subtraction schemes that have been developed over the last decades.

One class of schemes are the dipole subtraction methods, with the most prominent being CS subtraction [40, 85], implemented e.g. in Sherpa [86]. In these dipole schemes, the singular structure is in general represented by three indices, where each index stands for a particle – the emitter, the emitted, and the spectator. The latter is identical to the color partner in the eikonal approximation in the soft limit. Therefore, the dipole subtraction is formulated with an emphasis on soft radiation and features $3 \to 2$ mappings from Φ_{n+1} to the n-body phase space. Additionally, this involves summing over the three indices of the particles that represent each singular configuration, leading to a high number of counterterms. However, it allows for a highly generic parametrization of the $(n+1)$-body phase space. Besides CS, other variations of dipole subtraction schemes exist, like Nagy-Soper subtraction [87, 88] or Antenna subtraction [89–91]. The former is based on splitting functions in an improved parton shower [92–94] and reduces the number of subtraction terms [95]. Antenna subtraction shows promising developments towards NNLO automation (see [96], with many following). Both schemes have yet to be implemented for arbitrary processes in a multi-purpose MC event generator, however.

Besides the dipole approach, another scheme is FKS subtraction [41, 42, 66, 68], implemented e.g. in MG5_aMC@NLO [33], `POWHEG BOX` [59, 67], and WHIZARD [26, 97, 98]. In this scheme, the focus lies on collinear splittings, leading to singular configurations being represented by a pair of particles – the emitter and the emitted. Therefore, the number of indices that is summed over is reduced to two and we have a $2 \to 1$ mapping from an underlying n-body phase space to Φ_{n+1}. Consequently, as stated in [42], in general the number of counterterms is smaller than in CS subtraction and scales slower with increasing final-state multiplicity, benefiting numerical stability. As further listed in [42], FKS subtraction has many additional advantages, including better parallelization possibilities and the ease of supporting fully-polarized processes for future lepton colliders like the ILC. Another point given there is that because of its structure being akin to that of collinear radiation, it is straightforward to implement a

matching to parton-shower methods that are not based on color dipoles as e.g. provided by MC@NLO or POWHEG.

We do not cover the resonance-aware extension of FKS subtraction [69] in this thesis. Its implementation in WHIZARD has been detailed in [26] and applied to fully-differential top-pair production at a lepton collider in [99]. This formulation of FKS subtraction improves the behavior of the calculation for processes in which intermediate particle with very small decay widths are present. The contributing decay of these particles experiences real emissions that shift the virtuality of the real-emission propagator away from its value at Born level. This discrepancy leads to the mismatch in the cancellation of divergences and is treated in the resonance-aware FKS subtraction by additional terms and a modified phase-space parametrization.

We start presenting the theory behind FKS subtraction as established by [41, 42, 59, 66–69] by first discussing some general phase-space considerations in section 2.1, where we also work out the generalized notation based on [26, 42, 67] to be used throughout this thesis. In section 2.2 we then detail the construction of the real-subtraction terms for both FSR and ISR in QCD. Their integration over the radiation phase space and addition to the bare squared virtual amplitude $\mathcal{V}_b$ as in eq. (2.0.3) is laid out in section 2.3.

Before we continue, however, we comment on mixed NLO corrections in the section below.

Mixed Corrections at NLO

In this thesis we will focus on NLO QCD corrections. This does not cover the whole picture and needs some clarification, however, as generally in a gauge invariant computation, a process receives contributions from terms proportional to different orders of both QCD and EW couplings. Therefore, in general, we also need to expand the perturbation series in more than one coupling, leading to mixed corrections at higher orders. This procedure is explained in detail e.g. in [33]. We will simply present the example given therein of the partonic process $uu \to uu$, where only QCD and QED contributions are considered (i.e. no Z-boson is exchanged). The square of the sum of Born matrix elements gives terms proportional to the square of either the strong coupling α_S or the QED coupling α, as well as a mixed term proportional to $\alpha_S\alpha$. This is illustrated by the upper row of circles in fig. 2.1.

At NLO, each term of the squared Born amplitude receives both QCD and QED corrections due to the expansion in both α_S and α. In fig. 2.1 this is denoted by an arrow pointing to the left (QCD) and an arrow pointing to the right (QED). The resulting contributions are represented in the lower row of circles by their coupling orders. It is important to note that there the terms proportional to $\alpha_S^2\alpha$ and $\alpha_S\alpha^2$ each

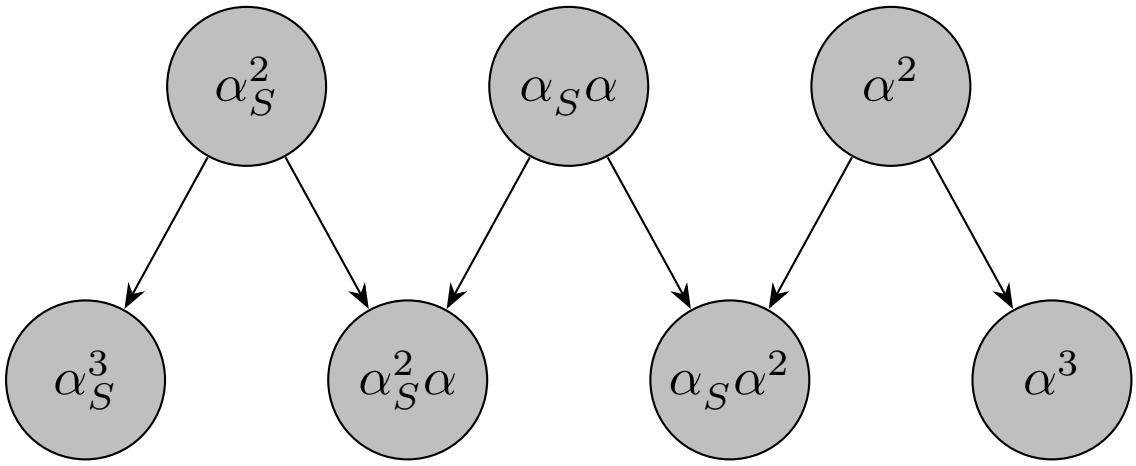


Figure 2.1. Illustration of the mixing of NLO QCD and QED corrections to the partonic process $uu \to uu$. The upper row shows the LO contributions proportional to combinations of different orders of the QCD and QED couplings α_S and α, respectively. The lower row represents their corresponding NLO corrections, with left arrows representing QCD corrections and right arrows those for QCD. The graphic is inspired by [33].

are the combined result of two types of corrections, each stemming from another Born contribution. This means that the divergences in these higher-order terms are only fully canceled when taking into account both QCD and QED corrections. Therefore, in principle, all types of corrections have to be taken into account for a physically meaningful computation of the full process at NLO. It is possible, however, to arrive at a physically sound result by only calculating the *pure* QCD (QED) corrections to the Born contribution with the highest order in α_S (α) since they do not receive mixed corrections. This is illustrated in fig. 2.1 by the leftmost (rightmost) circle in the lower row. In this thesis, we refer to the pure correction whenever we mention NLO QCD corrections.

2.1. Phase-space Generalities and Notation

In this thesis, we focus on NLO corrections at FO in the perturbation series. Consequently, we often use the terms *NLO* and *fNLO* interchangeably and only explicitly denote the FO prefix where we deem necessary.

We will denote the square of a four-momentum p of a particle with invariant mass m as

$$p^2 = \left(p^0\right) - |\mathbf{p}|^2 = m^2\,, \tag{2.1.1}$$

where we write the energy component of p as p^0 and its three-momentum as $\mathbf{p}$.

Deferring the handling of PDFs to section 2.2.2, we consider an arbitrary hard $2 \to n$ scattering process and write the corresponding set of Born four-momenta as

$$\Phi_n \equiv \{p_\oplus, p_\ominus, p_3, ..., p_{n+2}\}\,, \tag{2.1.2}$$

where the index $\oplus = 1$ and $\ominus = 2$ serve the purpose of denoting the direction of the incoming particles as parallel and anti-parallel to the beam axis direction $\oplus$, respectively. This will be especially handy once we consider PDFs. The set of four-momenta for a real emission of a gluon or the splitting of a gluon into a pair of partons is represented by

$$\Phi_{n+1} \equiv \{k_\oplus, k_\ominus, k_3, ..., k_{n+3}\}\,. \tag{2.1.3}$$

Both sets must satisfy momentum conservation, respectively given by

$$p_\oplus + p_\ominus = \sum_{i=3}^{n+2} p_i\,, \quad \text{and} \quad k_\oplus + k_\ominus = \sum_{i=3}^{n+3} k_i\,. \tag{2.1.4}$$

Furthermore, we define

$$s \equiv (p_\oplus + p_\ominus)^2 \tag{2.1.5}$$

and write the n-body phase space element as

$$d\Phi_n\left(p_\oplus + p_\ominus; p_3, ..., p_{n+2}\right) = (2\pi)^4\,\delta^4\left(p_\oplus + p_\ominus - \sum_{i=3}^{n+2} p_i\right)\prod_{i=3}^{n+2}\frac{d^3p_i}{(2\pi)^3\,2p_i^0}\,. \tag{2.1.6}$$

Phase-space singularities in the computation of n-body (or Born-like) cross sections are assumed to be taken care of through the adequate choice of kinematic cuts applied to the final-state particles. We adapt the notation from [42] and write this condition as

$$J^{n_L^{(B)}}\,. \tag{2.1.7}$$

For QCD, this condition can generally be met by demanding that all the strongly-interacting light particles $n_L^{(B)}$ of the Born-like final state are resolvable individually as a jet by an arbitrary IR-safe jet algorithm. Any remaining QED singularities are understood to be dealt with in eq. (2.1.7) as well.

The real phase space, however, retains IR singularities which have to be treated with a subtraction formalism in an MC calculation as stated in the introduction of chapter 2. For each of these singular phase-space regions we can define a mapping $\mathrm{M}^{(\alpha)}$ as explained in [67] such that

$$\tilde{\Phi}_{n+1}^{(\alpha)} \equiv \mathrm{M}^{(\alpha)}\left(\Phi_{n+1}\right) = \left\{\tilde{k}_\oplus^{(\alpha)}, \tilde{k}_\ominus^{(\alpha)}, \tilde{k}_3^{(\alpha)}, \cdots, \tilde{k}_{n+3}^{(\alpha)}\right\} \tag{2.1.8}$$

is a set of four-vectors containing a singular configuration. We require each mapping $\mathrm{M}^{(\alpha)}$ to be the identity at its particular singular phase-space configuration and to be smooth in its vicinity. We will call this resulting part of phase space the α-region. While the tilde notation serves a purpose of disambiguation in this instance, we will later drop the notation of eq. (2.1.8) and simply label momenta belonging to a $(n+1)$ kinematics as k, since the context should always be clear.

In the FKS formalism, the singularities in a given α-region are associated to a pair of particle indices (i, j) in the real flavor structure f_r with kinematics that leads to collinear and/or soft divergences. To be more precise, a real flavor structure

$$f_r \equiv \{\mathcal{I}_1, \mathcal{I}_2, \cdots\} \tag{2.1.9}$$

is the set of particle flavors $\mathcal{I}_i$ corresponding to their respective momenta in a kinematic set $\tilde{\Phi}^{(\alpha)}_{n+1}$. Analogously, f_B is the Born flavor structure. We will also refer to f_r and f_B as the real and Born (sub)process, respectively.

Writing the particle indices as partons for the sake of general demonstration, the following types of QCD FKS pairs exist [26]:

1. (q, g) and $(\bar{q}, g)$, describing a gluon with vanishing energy emitted by an (anti-) quark, inducing a soft singularity, and/or a spatial momentum (anti-)parallel to the emitting parton, resulting in a collinear singularity if $m_{q/\bar{q}} = 0$.

2. $(q, \bar{q})$, describing a $g \to q\bar{q}$ splitting which induces a collinear singularity if both $m_{q/\bar{q}} = 0$. As we will later see explicitly, these splittings do not have a soft singularity.

3. (g, g), describing a $g \to gg$ splitting, giving rise to both soft and collinear divergences.

Generally, for the first two kinds of FKS pairs also the mirrored configurations (g, q), $(g, \bar{q})$, and $(\bar{q}, q)$ exist. These, however, have the same matrix elements as their mirrored counterpart and would contribute twice. Leaving these pairs out of the calculation eliminates the need for a symmetry factor and lets us define parton $\mathcal{I}_j$ to always be massless and thereby ensuring that each FKS pair induces a divergence. The radiation of a massive particle does not lead to a singularity since in this case it is regularized by the mass of the emitted particle, with the same being true for a massive emitter in a collinear splitting. Only in the case of (g, g) associated with FSR we keep the trivial mirror pair and apply a symmetry factor, described in section 2.1.2.

In order to more accurately define the FKS pairs, we adopt the labeling convention

for particles from [42] with the modifications from [26] and write

n_I as the number of initial-state particles,

n_F as the index of the first strongly-interacting particle, where

$$n_F = \begin{cases} 1 & \text{for hadron-hadron collisions,} \\ 2 & \text{for lepton-hadron collisions,} \\ 3 & \text{for lepton-lepton collisions,} \end{cases} \tag{2.1.10}$$

$n_L^{(B,R)}$ as the number of strongly-interacting light final-state particles (Born, Real),

n_H as the number of strongly-interacting heavy final-state particles,

$n_\emptyset$ as the number of final-state particles not interacting strongly,

n_S as the total number of all strongly interacting particles.

We then impose an ordering on the particle flavors $\mathcal{I}_i$ in flavor structures, both real and Born-like, as follows:

$$\begin{aligned}
&\text{initial state} && : \left\{ \mathcal{I}_i \,\middle|\, 1 \le i \le n_I \right\}, \\
&\text{final-state light partons} && : \left\{ \mathcal{I}_i \,\middle|\, n_I + 1 \le i \le n_I + n_L^{(B,R)} \right\}, \\
&\text{final-state heavy partons} && : \left\{ \mathcal{I}_i \,\middle|\, n_I + 1 + n_L^{(B,R)} \le i \le n_I + n_L^{(B,R)} + n_H \right\}, \\
&\text{non-QCD particles} && : \left\{ \mathcal{I}_i \,\middle|\, n_I + 1 + n_L^{(B,R)} + n_H \le i \le n_I + n_L^{(B,R)} + n_H + n_\emptyset \right\}.
\end{aligned}$$

With this labeling in place, we can now define the set of all FKS pairs associated with inducing a divergence in the squared matrix element $\mathcal{R}(f_r)$ belonging to a real flavor structure f_r as [26, 42]

$$\begin{aligned}
\alpha \in \mathcal{P}_{\text{FKS}}(f_r) = \Big\{ (i,j) \,\Big|\, & n_F \le i \le n_L^{(R)} + n_H + n_I,\ n_I + 1 \le j \le n_L^{(R)} + n_I, i \ne j, \\
& \mathcal{R}(f_r) J^{n_L^{(B)}} \to \infty \quad \text{if} \quad k_j^0 \to 0 \quad \text{and/or} \quad \mathbf{k}_i \parallel \mathbf{k}_j, \\
& \text{non-redundancy conditions} \Big\}.
\end{aligned} \tag{2.1.11}$$

While this definition allows $\mathcal{I}_i$ to be part of the initial state, it restricts $\mathcal{I}_j$ to the final-state. Therefore, we call $\mathcal{I}_j$ the *emitted*, or *FKS parton*. The non-redundancy conditions avoid the double counting of symmetric FKS pairs, with only FKS pairs with two gluons that are associated with FSR being kept symmetrical, as stated earlier.

They are defined as

$$\mathcal{I}_i \neq g, \mathcal{I}_j = g, \quad (i,j) \in \mathcal{P}_{\text{FKS}} \quad \Rightarrow \quad (j,i) \notin \mathcal{P}_{\text{FKS}} \quad \text{if} \quad n_I + 1 \leq i\,, \tag{2.1.12}$$

$$\mathcal{I}_i \neq g, \mathcal{I}_j \neq g, \quad (i,j) \in \mathcal{P}_{\text{FKS}} \quad \Rightarrow \quad (j,i) \notin \mathcal{P}_{\text{FKS}} \quad \text{if} \quad n_I + 1 \leq i < j\,. \tag{2.1.13}$$

We will use the notations α and (i,j) (or ij in case of a subscript) interchangeably, since they are unambiguously tied together. Depending on the context and the intended emphasis on either the α-region or its associated FKS pair, however, it is convenient to choose between the different notations.

Coming back to eq. (2.1.8), we can now associate with each singular set $\tilde{\Phi}_{n+1}^{(\alpha)}$ a Born-like set $\bar{\Phi}_n^{(\alpha)}$, also called the *underlying Born* kinematics. It is defined in the barred notation of [67] as

$$\bar{\Phi}_n^{(\alpha)} \equiv \left\{ \bar{k}_\oplus^{(\alpha)}, \bar{k}_\ominus^{(\alpha)}, \bar{k}_3^{(\alpha)}, \cdots, \bar{k}_{n+2}^{(\alpha)} \right\}. \tag{2.1.14}$$

This set of four-momenta can easily be constructed from $\tilde{\Phi}_{n+1}^{(\alpha)}$ by simply deleting the parton with zero momentum if α is associated with a soft divergence or – in case of a final-state collinear singularity – by replacing the pair of collinear momenta by their sum. For initial-state collinear configurations the momentum of the emitter is replaced with its momentum after emitting the collinear parton, whose momentum is then deleted from the list of external momenta. We then relabel the momenta in $\bar{\Phi}_n^{(\alpha)}$ according to the ordering introduced before. By constructing $\bar{\Phi}_n^{(\alpha)}$ in this way we also assure momentum conservation.

For each α-region, we can now parametrize the $(n+1)$-body kinematics as [67]

$$\Phi_{n+1} \overset{(\alpha)}{\Longleftrightarrow} \left\{ \bar{\Phi}_n^{(\alpha)}, \Phi_{\text{rad}}^{(\alpha)} \right\} \tag{2.1.15}$$

via the underlying Born configuration and a set $\Phi_{\text{rad}}^{(\alpha)}$ of variables describing the additional three degrees of freedom of the FKS parton. The bidirectional arrow indicates that the mapping in eq. (2.1.8) is invertible. Additionally, we note that for example for different α-regions, the configurations

$$\bar{\Phi}_n^{(\alpha)} = \bar{\Phi}_n^{(\alpha')} = \bar{\Phi}_n^{(\alpha'')} = \cdots\,, \quad \Phi_{\text{rad}}^{(\alpha)} \neq \Phi_{\text{rad}}^{(\alpha')} \neq \Phi_{\text{rad}}^{(\alpha'')} \neq \cdots \tag{2.1.16}$$

are allowed as long as all α-regions have the same underlying Born flavor structure. Therefore, the same underlying Born configuration can be associated with different real configurations, whereas the other way around a single real configuration can only

have one underlying Born configuration.[1]

Equation (2.1.15) allows us to factorize the $(n+1)$-body phase-space element as

$$d\Phi_{n+1}^{(\alpha)} = d\bar{\Phi}_n^{(\alpha)} d\Phi_{\text{rad}}^{(\alpha)} \,. \tag{2.1.17}$$

Further, the radiation phase-space element $d\Phi_{\text{rad}}^{(\alpha)}$ can be written as

$$d\Phi_{\text{rad}}^{(\alpha)} = \mathcal{J}^{(\alpha)}(\xi_j, y_{ij}, \phi_j) d\xi_j dy_{ij} d\phi_j \,, \quad (i,j) \in \mathcal{P}_{\text{FKS}} \,, \tag{2.1.18}$$

in terms of a Jacobian $\mathcal{J}^{(\alpha)}$, with its explicit form given in appendix A.3. In general, $\mathcal{J}^{(\alpha)}$ is dependent both on $\bar{\Phi}_n^{(\alpha)}$ and the typical choice for the set of FKS variables $\{\xi, y, \phi\}$ that parametrize $\Phi_{\text{rad}}^{(\alpha)}$. Considering only massless emitting particles, in the center-of-mass (CM) frame of the incoming partons these variables are [41, 67]

- the rescaled energy of the emitted particle $\mathcal{I}_j$, given by

$$\xi_j = \frac{2k_j^0}{\sqrt{s}} \in [0, \xi_{\text{max}}] \,, \tag{2.1.19}$$

  where in the case of FSR the upper bound of ξ_{max} is limited by the energy $\bar{k}_i^0$ the emitting particle $\mathcal{I}_i$ has in the underlying Born kinematics such that

$$\xi_{\text{max}} = \frac{2\bar{k}_i^0}{\sqrt{s}} \leq 1 \,, \tag{2.1.20}$$

- the cosine between the spatial vectors of the emitted particle $\mathcal{I}_j$ and a reference direction, defined by either the incoming parton parallel to the $\oplus$ direction of the beam axis (for ISR) or its splitting partner $\mathcal{I}_i$ (for FSR), denoted by

$$[-1,1] \ni y_{ij} \equiv \frac{\mathbf{k}_i \cdot \mathbf{k}_j}{|\mathbf{k}_i|\,|\mathbf{k}_j|} = \cos\theta_{ij} \quad \text{for FSR,} \tag{2.1.21}$$

$$\text{and} \quad [-1,1] \ni y_j \equiv \frac{\mathbf{k}_\oplus \cdot \mathbf{k}_j}{|\mathbf{k}_\oplus|\,|\mathbf{k}_j|} = \cos\theta_{\oplus j} \quad \text{for ISR,} \tag{2.1.22}$$

- and the azimuthal angle of the emitted particle around the same reference direction written as

$$\phi_j \in [0, 2\pi] \,. \tag{2.1.23}$$

We suppress the indices of the FKS variables for the ease of notation in general cases or where the context is clear. While the definition of ϕ always remains the same, the

[1] This is in contrast to CS subtraction where there can be more than one Born configuration for a single real one.

expressions for $\xi_{\max}$ and y can become more complex, e.g. for massive emitters or ISR, as described in appendices A.1.1 and A.2, respectively. In terms of the FKS variables, a soft singularity corresponds to $\xi \to 0$, whereas $y \to \pm 1$ induces a collinear singularity. The negative sign of y corresponds to the case of ISR where the radiated particle is emitted by and collinear to the incoming particle in $\ominus$ direction, therefore becoming anti-collinear to the beam direction $\oplus$.

2.1.1. Real Phase-space Generation

The actual construction of the $(n+1)$-body phase space, together with the explicit forms of the Jacobians $\mathcal{J}^{(\alpha)}$, will be deferred to appendix A. Although the correct generation of real-emission kinematics is imperative for assuring sufficient cancellation of divergences by the subtraction counterterms, its details are not relevant for the discussions that follow.

In general, the idea of constructing the real-emission phase space in a MC event generator is rather straightforward. Since the arrow in eq. (2.1.15) indicates that the correspondence is one to one and the mapping in eq. (2.1.8) is invertible, we can devise an inverse construction of Φ_{n+1}, given a set of barred kinematics $\bar{\Phi}_n^{(\alpha)}$ and a set of FKS variables $\Phi_{\text{rad}}^{(\alpha)}$. Considering the fact that $\bar{\Phi}_n^{(\alpha)}$ is for all practical means just a regular n-body phase space for the flavor structure of the particular underlying Born associated with one or more real flavor structures, any MC event generator should be able to provide this part already. The additional steps are then to generate an extra set of random numbers for the FKS variables and use them, together with the n-body kinematics, to calculate the kinematics for the $(n+1)$-body phase space, as outlined in appendix A.

2.1.2. Partitioning the Real Phase Space

The Partition Functions $\mathcal{S}$

From what we have seen in section 2.1, in the FKS subtraction scheme the $(n+1)$-body phase space can be effectively partitioned into regions containing at most one soft and one collinear singularity associated with an FKS pair $(i,j) \in \mathcal{P}_{\text{FKS}}$. We can use this to split up the squared amplitude $\mathcal{R}(f_r)$ of the real matrix element, belonging to a flavor structure f_r of a real-emission (sub)process, into a sum of terms [41, 42, 67],

$$\mathcal{R}(f_r) = \sum_{(i,j)\in\mathcal{P}_{\text{FKS}}} \mathcal{S}_{ij}(f_r)\mathcal{R}(f_r) \quad \text{with} \quad \mathcal{S}_{ij}(f_r)\mathcal{R}(f_r) \equiv \mathcal{R}_{ij}(f_r)\,, \tag{2.1.24}$$

so that $\mathcal{R}_{ij}(f_r)$ will diverge only if the FKS parton $\mathcal{I}_j$ becomes soft and/or collinear to $\mathcal{I}_i$ while remaining finite everywhere else. For ease of notation we will understand the dependence on f_r as implied and only write it explicitly when needed.

There is some freedom in choosing the explicit form of the $\mathcal{S}$-functions. However, some important constraints have to be met, the most obvious being

$$\sum_{(i,j)\in\mathcal{P}_{\text{FKS}}} \mathcal{S}_{ij} = 1\,. \tag{2.1.25}$$

Additionally, we require $\mathcal{S}_{ij}$ to vanish in all singular regions of phase space that are associated with FKS pairs different from (i,j) [42]:

$$\lim_{\mathbf{k}_k\parallel\mathbf{k}_l} \mathcal{S}_{ij} = 0 \quad \forall\,\{k,l\} \neq \{i,j\} \mid (k,l) \in \mathcal{P}_{\text{FKS}}\,, \tag{2.1.26}$$

$$\lim_{k_k^0\to 0} \mathcal{S}_{ij} = 0 \quad \forall k \mid \mathcal{I}_k = g \quad \text{and} \quad \exists l \mid ((l,k) \in \mathcal{P}_{\text{FKS}} \vee (k,l) \in \mathcal{P}_{\text{FKS}})\,. \tag{2.1.27}$$

We recall that collinear divergences are regulated when at least one splitting partner is massive. In these cases the $\mathcal{S}$-functions behave as they do in any other regular phase-space region by taking a value between 0 and 1. In a collinear region associated with an FKS pair having two massless members, however, we demand [42, 67]

$$\lim_{\mathbf{k}_i\parallel\mathbf{k}_j} \mathcal{S}_{ij} = h_{ij}(z_{ij}) \quad \text{with} \quad 0 \leq z_{ij} \leq 1 \quad \text{and} \quad z_{ij} = \frac{k_j^0}{k_i^0 + k_j^0}\,. \tag{2.1.28}$$

The functions h_{ij} only differ from unity if both $\mathcal{S}_{ij}$ and $\mathcal{S}_{ji}$ are considered, which we only do in the case of FSR when both members of an FKS pair are gluons, as defined in eq. (2.1.11). Therefore, we have [26, 67]

$$h_{ij}(z_{ij}) = \begin{cases} h(z_{ij}) \geq 0 & \text{if} \quad (\mathcal{I}_i,\mathcal{I}_j) = (g,g)\,, \quad i > n_I \\ 1 & \text{else} \end{cases} \tag{2.1.29}$$

and require [42, 67]

$$\lim_{z_{ij}\to 0} h(z_{ij}) = 1\,, \quad \lim_{z_{ij}\to 1} h(z_{ij}) = 0\,, \quad h(z_{ij}) + h(1-z_{ij}) = 1\,. \tag{2.1.30}$$

A typical choice for $h(z_{ij})$ is [67]

$$h(z_{ij}) = \frac{(1-z_{ij})^a}{z_{ij}^a + (1-z_{ij})^a}\,, \quad 0 < a \in \mathbb{R}\,. \tag{2.1.31}$$

Together with the conditions in eq. (2.1.30), this choice prevents the double-counting

of $\mathcal{S}_{ij}$ and $\mathcal{S}_{ji}$. The exponent a can be chosen freely, as the numeric results should not be influenced by a. In the region of phase space where the emitted parton $\mathcal{I}_j$ induces a soft singularity we require [42, 67]

$$\lim_{k_j^0 \to 0} \mathcal{S}_{ij} = c_{ij} \quad \forall j \mid \mathcal{I}_j = g \,, \tag{2.1.32}$$

where $c_{ij} \in \mathbb{R}$ such that

$$0 < c_{ij} \leq 1 \quad \text{and} \quad \sum_{\substack{i \\ (i,j) \in \mathcal{P}_{\text{FKS}}}} c_{ij} = 1 \,. \tag{2.1.33}$$

Choice of $\mathcal{S}$-functions

The constraints on the $\mathcal{S}$-functions as given in eqs. (2.1.25) to (2.1.33) still leave the freedom to choose a specific form for them. It was shown in [100] in the context of NLO QCD single-top hadroproduction with `MC@NLO` that smooth $\mathcal{S}$-functions improve the stability of the numeric integration when compared to the original representation in terms of products of Heaviside step functions [41, 66]. The prevalent choice of $\mathcal{S}$-functions in modern FKS literature [42, 67] is

$$\mathcal{S}_{ij} = \frac{1}{\mathcal{D} d_{ij}} h_{ij}(z_{ij}) \,, \quad (i,j) \in \mathcal{P}_{\text{FKS}} \,, \tag{2.1.34}$$

$$\mathcal{D} \equiv \sum_{(k,l) \in \mathcal{P}_{\text{FKS}}} \frac{1}{d_{kl}} h_{kl}(z_{kl}) \,, \tag{2.1.35}$$

with $h_{ij}(z_{ij})$ as defined before. The phase-space weights d_{ij} are required to fulfill the conditions

$$d_{ij} = 0 \quad \text{if and only if} \quad k_i^0 = 0 \quad \text{or} \quad k_j^0 = 0 \quad \text{or} \quad \mathbf{k}_i \parallel \mathbf{k}_j \tag{2.1.36}$$

and can in general be chosen as [59, 67]

$$d_{ij} = \left[2 k_i \cdot k_j \frac{k_i^0 k_j^0}{\left(k_i^0 + k_j^0\right)^2} \right]^p \underset{m_i \to 0}{\overset{m_j \to 0}{=}} 2^p \left(\frac{k_i^0 k_j^0}{k_i^0 + k_j^0} \right)^{2p} (1 \mp y)^p \,, \tag{2.1.37}$$

with an arbitrary exponent $0 < p \in \mathbb{R}$, which we will set to one for convenience. The plus in front of the y in the last bracket covers the case for ISR, where additionally $y = -1$ induces a divergence when the emitted parton becomes collinear to the incoming parton in $\ominus$-direction. In this case, d_{ij} must vanish as well.

In some cases we can, and even have to, simplify d_{ij}. It is required in the limit in which the FKS pair (i,j) induces a soft singularity. By rewriting the denominator of

eq. (2.1.34) using the behavior of h_{ij} defined in eqs. (2.1.29) and (2.1.30) such that[2]

$$\mathcal{D}d_{ij} = 1 + \sum_{\substack{k|(k,j)\in\mathcal{P}_{\mathrm{FKS}} \\ k\neq i}} \frac{d_{ij}}{d_{kj}} + \sum_{\substack{(k,l)\in\mathcal{P}_{\mathrm{FKS}} \\ l\neq j}} \frac{d_{ij}}{d_{kl}}\,, \tag{2.1.38}$$

we can see that in the soft limit, where both d_{ij} and d_{kj} vanish, the ratio between the phase-space weights in the second term of eq. (2.1.38) becomes ill-defined. However, we can see that in the soft limit the ratio becomes independent of the radiated energy k_j^0 by writing $k_j = k_j^0 \hat{k}_j$ and using eq. (2.1.37), which gives

$$\lim_{k_j^0\to 0} \frac{d_{ij}}{d_{kj}} = \lim_{k_j^0\to 0} \frac{2\left(k_i\cdot\hat{k}_j\right) k_i^0 \left(k_j^0\right)^2 \left(k_k^0+k_j^0\right)^2}{2\left(k_k\cdot\hat{k}_j\right) k_k^0 \left(k_j^0\right)^2 \left(k_i^0+k_j^0\right)^2} = \frac{2\left(k_i\cdot\hat{k}_j\right) k_k^0}{2\left(k_k\cdot\hat{k}_j\right) k_i^0}\,. \tag{2.1.39}$$

Therefore, we have to use the simplified phase-space weights in the soft limit, given by

$$d_{ij}^{\mathrm{soft}} = \frac{2k_i\cdot\hat{k}_j}{k_i^0}\,. \tag{2.1.40}$$

We note that because of the definition of FKS pairs in eq. (2.1.11) the emitted FKS parton is always massless, making $\hat{\mathbf{k}}_j$ a unit vector with the direction of the emitted particle and $\hat{k}_j^0 = 1$. Also, in the soft limit the real momentum k_i becomes equal to its underlying Born momentum $\bar{k}_i$.

For the case of ISR, we choose to simplify the d_{ij} functions by evaluating the second equality of eq. (2.1.37) in the limit $k_i^0 \gg k_j^0$, where the energy sum in the denominator effectively cancels the k_i^0 in the numerator. This is a valid choice for the initial-state d_{ij} functions, since we have chosen not to symmetrize any FKS pairs associated to ISR divergences. Hence, we do not have to worry about $k_{1/2}^0 \to 0$, while still meeting the conditions for d_{ij} and $\mathcal{S}_{ij}$. Demanding any incoming QCD partons to be massless, we can then write

$$d_{1j} = 2\left(k_j^0\right)^2 (1-y_j)\,, \tag{2.1.41}$$

$$d_{2j} = 2\left(k_j^0\right)^2 (1+y_j)\,, \tag{2.1.42}$$

$$d_{0j} = \left(k_j^0\right)^2 \left(1-y_j^2\right)\,, \quad d_{0j}^{-1} = d_{1j}^{-1} + d_{2j}^{-1}\,. \tag{2.1.43}$$

[2] If $h_{ij} \neq 1$, i.e. both members of an FKS pair being gluons, we use eq. (2.1.30) to obtain:

$$\frac{1}{d_{ij}} h_{ij}(z_{ij}) + \frac{1}{d_{ji}} h_{ji}(z_{ji}) = \frac{1}{d_{ij}} \left(h_{ij}(z_{ij}) + h_{ij}(1-z_{ij})\right) = \frac{1}{d_{ij}}$$

Here, the notation of an emitter with index $i = 0$ is adapted from [59]. This notation simply combines α-regions associated with collinear divergences stemming from both the $\oplus$ and $\ominus$ directions that do not change the underlying Born kinematics (and flavors). According to the definition of $\mathcal{P}_{\text{FKS}}$ in eq. (2.1.11) and the non-redundancy conditions eqs. (2.1.12) and (2.1.13), this case can only occur for the initial-state real emission of a gluon. Again, in the ISR soft limit one can easily show that by replacing d_{kj} in eq. (2.1.39) with the ones from eqs. (2.1.41) to (2.1.43), respectively, we obtain

$$d_{1j}^{\text{soft}} = 2\,(1 - y_j)\;, \tag{2.1.44}$$

$$d_{2j}^{\text{soft}} = 2\,(1 + y_j)\;, \tag{2.1.45}$$

$$d_{0j}^{\text{soft}} = 1 - y_j^2\,. \tag{2.1.46}$$

2.2. Contributions to the Real Cross Section

With the notation and the phase-space generalities discussed, we draw our attention to the construction of the real-subtraction terms. In general, the total NLO real cross section for a real $2 \to n+1$ (sub)process f_r can be written as the sum of integrals over the partitioned phase space and the corresponding partitioned squared real amplitude such that

$$\sigma_{\text{NLO}}^{\text{real}}(f_r) = \frac{J^{n_L^{(B)}}}{\mathcal{N}(f_r)} \int d\Phi_{n+1} \mathcal{R}(f_r) = \frac{J^{n_L^{(B)}}}{\mathcal{N}(f_r)} \sum_{\alpha \in \mathcal{P}_{\text{FKS}}} \int d\bar{\Phi}_n^{(\alpha)} d\Phi_{\text{rad}}^{(\alpha)} \mathcal{R}_\alpha(f_r)\,. \tag{2.2.1}$$

Here, we explicitly wrote out the condition $J^{n_L^{(B)}}$ as stated in eq. (2.1.7) and the symmetry factor $\mathcal{N}(f_r)$ that accounts for the presence of identical particles in the final state of the real (sub)process f_r. We can recall from eq. (1.0.2) that the divergences in eq. (2.2.1) stem from propagators proportional to $[\xi\,(1-y)]^{-1}$ (for FSR) in terms of the FKS variables. Applying dimensional regularization makes the divergences explicit and thus we can write the corresponding phase space in $d = 4 - 2\epsilon$ dimensions for each term of the sum as [41, 67]

$$d\bar{\Phi}_n^{(\alpha)} d\Phi_{\text{rad}}^{(\alpha)} = (2\pi)^d\, \delta^d \left(k_\oplus + k_\ominus - \sum_{l=3}^{n+3} k_l \right) \left[\prod_{\substack{l=3 \\ l \neq j}}^{n+3} \frac{d^{d-1} k_l}{(2\pi)^{d-1}\, 2k_l^0} \right]$$

$$\times \underbrace{\frac{s^{1-\epsilon}}{(4\pi)^{d-1}} \xi_j^{1-2\epsilon} \left(1 - y^2\right)^{-\epsilon} d\xi_j dy d\Omega^{d-2}}_{d\Phi_{\text{rad}}^{(\alpha)}}\,, \tag{2.2.2}$$

where we denote the quantities of the emitted particle with index j as usual. The azimuthal integral over the whole range of ϕ in $d = 4 - 2\epsilon$ dimensions gives [101]

$$\int d\Omega^{d-2} = \frac{2\pi^{1-\epsilon}}{\Gamma(1-\epsilon)}\,. \tag{2.2.3}$$

In the following, we will first consider final-state singularities to establish the technicalities needed in constructing the real counterterms. We will return to the more involved treatment of ISR in section 2.2.2.

2.2.1. Final-state Subtraction Terms

Final-state collinear divergences can only be induced by $y_{ij} \to 1$. Hence, we can define a regularized squared amplitude $\tilde{\mathcal{R}}_{ij}$ such that [41, 67]

$$\mathcal{R}_\alpha = \frac{1}{\xi_j^2}\frac{1}{1-y_{ij}}\tilde{\mathcal{R}}_{ij}(\xi_j, y_{ij})\,, \quad \tilde{\mathcal{R}}_{ij}(\xi_j, y_{ij}) \equiv \xi_j^2\,(1-y_{ij})\,\mathcal{R}_\alpha\,, \quad (i,j) \in \mathcal{P}_{\text{FKS}}\,. \tag{2.2.4}$$

For one FKS pair the integration over the radiation phase space in eq. (2.2.1) is then proportional to

$$\int d\Phi_{\text{rad}}^{(\alpha)}\mathcal{R}_\alpha(f_r) \sim \int_0^{\xi_{\max}} d\xi_j \xi_j^{-1-2\epsilon}\int_{-1}^{1} dy_{ij}\frac{(1-y_{ij})^{-1-\epsilon}}{(1+y_{ij})^{\epsilon}}\int d\Omega^{d-2}\tilde{\mathcal{R}}_{ij}(\xi_j, y_{ij})\,. \tag{2.2.5}$$

We will deal with the divergences that occur for $\epsilon = 0$ with the help of modified *plus distributions*, explained in the following excursion.

Modified Plus Distributions

Suppose we have an integral I of the form

$$I \equiv \int_0^{\xi_{\max}} d\xi \xi^{-1-2\epsilon} f(\xi, y)\,, \tag{2.2.6}$$

with $f(\xi, y)$ being a finite function in the integration volume. The integral then has an obvious singularity for $\xi = 0$ if $\epsilon = 0$, which we isolate in a separate term by writing

$$I = \int_0^{\xi_{\max}} d\xi \frac{f(\xi,y) - f(0,y)\Theta(\xi_{\text{c}} - \xi)}{\xi^{1+2\epsilon}} + f(0,y)\int_0^{\xi_{\max}} d\xi \xi^{-1-2\epsilon}\Theta(\xi_{\text{c}} - \xi)\,, \tag{2.2.7}$$

where $0 < \xi_{\text{c}} \leq \xi_{\max}$ is an arbitrary parameter on which the final result should not depend. Writing $\xi^{-2\epsilon} = e^{-2\epsilon\log\xi}$ as a Taylor series and integrating the second

term of eq. (2.2.7), we get

$$I = \int_0^{\xi_{\max}} d\xi \frac{f(\xi,y) - f(0,y)\Theta(\xi_c - \xi)}{\xi} \left[\sum_{n=0}^{\infty} \frac{(-2\epsilon)^n}{n!} \log^n \xi\right] - \frac{\xi_c^{-2\epsilon}}{2\epsilon} f(0,y) \tag{2.2.8}$$

$$= \int_0^{\xi_{\max}} d\xi \left\{ -\frac{\xi_c^{-2\epsilon}}{2\epsilon}\delta(\xi) + \sum_{n=0}^{\infty} \frac{(-2\epsilon)^n}{n!} \left(\frac{\log^n \xi}{\xi}\right)_c \right\} f(\xi,y)\,, \tag{2.2.9}$$

where we have defined the modified plus distributions denoted with the subscript c in the last step. Analogously, we can repeat the steps above for the expression $(1 \mp y)^{-1-\epsilon}$ in eq. (2.2.5) and thus get

$$\xi^{-1-2\epsilon} = -\frac{\xi_c^{-2\epsilon}}{2\epsilon}\delta(\xi) + \left(\frac{1}{\xi}\right)_c - 2\epsilon \left(\frac{\log \xi}{\xi}\right)_c + \mathcal{O}(\epsilon^2)\,, \tag{2.2.10}$$

$$(1 \mp y)^{-1-\epsilon} = -\frac{\left(\delta_{I/O}\right)^{-\epsilon}}{\epsilon}\delta(1 \mp y) + \left(\frac{1}{1 \mp y}\right)_{\delta_{I/O}} + \mathcal{O}(\epsilon)\,, \tag{2.2.11}$$

where $0 < \delta_{I/O} \leq 2$ are arbitrary parameters on which the final result should not depend. The two indices I and O will be used to indicate that we can have different values for initial and final state. For a quick overview we state the definition of the action of the distributions used within this thesis explicitly:

$$\int_0^{\xi_{\max}} d\xi \left(\frac{1}{\xi}\right)_c f(\xi,y) = \int_0^{\xi_{\max}} d\xi \frac{f(\xi,y) - f(0,y)\Theta(\xi_c - \xi)}{\xi}\,, \tag{2.2.12}$$

$$\int_0^{\xi_{\max}} d\xi \left(\frac{\log \xi}{\xi}\right)_c f(\xi,y) = \int_0^{\xi_{\max}} d\xi \frac{f(\xi,y) - f(0,y)\Theta(\xi_c - \xi)}{\xi} \log \xi\,, \tag{2.2.13}$$

$$\int_{-1}^{1} dy \left(\frac{1}{1 \mp y}\right)_{\delta_{I/O}} f(\xi,y) = \int_{-1}^{1} dy \frac{f(\xi,y) - f(\xi,\pm 1)\Theta(\pm y - 1 + \delta_{I/O})}{1 \mp y}\,. \tag{2.2.14}$$

The more common definition of the unmodified plus distributions is such that

$$\left(\frac{\log^n \xi}{\xi}\right)_+ = \left(\frac{\log^n \xi}{\xi}\right)_c \quad \text{for} \quad \xi_c = \xi_{\max} = 1\,, \tag{2.2.15}$$

$$\left(\frac{1}{1 \mp y}\right)_+ = \left(\frac{1}{1 \mp y}\right)_{\delta_{I/O}} \quad \text{for} \quad \delta_{I/O} = 2\,. \tag{2.2.16}$$

Returning to eq. (2.2.5), we can now use eqs. (2.2.10) and (2.2.11) in order to write

$$\begin{aligned}\int d\Phi_{\text{rad}}^{(\alpha)}\mathcal{R}_\alpha(f_r) = \frac{s^{1-\epsilon}}{(4\pi)^{d-1}}\Bigg\{ &\int d\Omega^{d-2}\frac{\xi_{\text{c}}^{-2\epsilon}\,(2\delta_O)^{-\epsilon}}{2\epsilon^2}\tilde{R}_{ij}(0,1) + \mathcal{O}(\epsilon) \qquad (2.2.17)\\ &- \int d\Omega^{d-2}\int_{-1}^{1} dy_{ij}\,(1+y_{ij})^{-\epsilon}\,\frac{\xi_{\text{c}}^{-2\epsilon}}{2\epsilon}\left[\left(\frac{1}{1-y}\right)_{\delta_O} + \mathcal{O}(\epsilon)\right]\tilde{\mathcal{R}}_{ij}(0,y_{ij})\\ &- \int d\Omega^{d-2}\int_0^{\xi_{\max}} d\xi_j\frac{(2\delta_O)^{-\epsilon}}{\epsilon}\left[\left(\frac{1}{\xi_j}\right)_{\text{c}} - 2\epsilon\left(\frac{\log\xi_j}{\xi_j}\right)_{\text{c}}\right]\tilde{\mathcal{R}}_{ij}(\xi_j,1)\Bigg\}\\ &+ \int d\phi_j\int_0^{\xi_{\max}} d\xi_j\int_{-1}^{1} dy_{ij}\left(\frac{1}{\xi_j}\right)_{\text{c}}\left(\frac{1}{1-y_{ij}}\right)_{\delta_O}\frac{\mathcal{J}^{(\alpha)}}{\xi_j}\tilde{\mathcal{R}}_{ij}(\xi_j,y_{ij})\\ &\equiv I_{s\delta,\alpha} + I_{s+,\alpha} + I_{+\delta,\alpha} + I_{++,\alpha}\,, \qquad (2.2.18)\end{aligned}$$

where we get a factor $2^{-\epsilon}$ in the first and third row due to $\delta(1-y_{ij})$ acting on $(1+y_{ij})^{-\epsilon}$. In the last line we defined a shorthand notation[3] for the integrals, where the subscripts s and δ stand for soft and collinear divergent contributions, while a $+$ denotes their regularization via the modified plus distributions, respectively.

Equation (2.2.17) leaves us with a very tameable set of terms, with the first three being explicit in their divergences. In detail, the first term ($I_{s\delta,\alpha}$) is soft-collinear divergent, while the second ($I_{s+,\alpha}$) and third ($I_{+\delta,\alpha}$) are soft and collinear divergent, respectively. According to the KLN theorem, however, these divergences will cancel against their respective terms in the squared one-loop matrix elements after the integration, thus giving rise to the counterterms for the virtual contribution.

The first two terms originate from the delta distribution in ξ, allowing us to calculate them by using the eikonal approximation for soft emissions in $d = 4-2\epsilon$ dimensions. The calculation of the integrated collinear term $I_{+\delta,\alpha}$ makes use of the fact that we can substitute the invariant amplitude with its collinear limit, eliminating the need to know $\tilde{\mathcal{R}}_{ij}$ in d dimensions. We will return to the integrated subtraction terms in section 2.3.

The last term ($I_{++,\alpha}$) of eq. (2.2.17) has no ϵ-poles left and the modified plus distributions act on the regularized squared amplitude $\tilde{\mathcal{R}}_{ij}$ as defined in eq. (2.2.4). Therefore, no singularities remain and we can safely return to four dimensions by setting $\epsilon = 0$. Consequently, the angular measure $d\Omega^{d-2}$ transforms into $d\phi_j$. In addition, here we explicitly write out the Jacobian $\mathcal{J}^{(\alpha)}(\xi_j, y_{ij}, \phi_j)$, which in general depends on the FKS variables and $\bar{\Phi}_n^{(\alpha)}$ and is given in detail in appendix A.3. We note that the Jacobian includes the ξ from eq. (2.2.2) used in the derivation of eq. (2.2.17), which we therefore have to cancel out again.

[3] While we borrow the style of the notation from [26, 69], we define the integrals I without the integration over the n-body phase space.

By applying the definitions of the modified plus distributions from eqs. (2.2.12) and (2.2.14) and defining

$$\hat{\mathcal{R}}_{ij}(\xi_j, y_{ij}) \equiv \frac{\mathcal{J}^{(\alpha)}(\xi_j, y_{ij}, \phi_j)}{\xi_j} \tilde{\mathcal{R}}_{ij}(\xi_j, y_{ij}) \tag{2.2.19}$$

we can write the last term of eq. (2.2.17) as [42]

$$\begin{aligned} I_{++,\alpha} = \int_0^{2\pi} d\phi_j \int_0^{\xi_{\max}} \frac{d\xi_j}{\xi_j} \int_{-1}^{1} \frac{dy_{ij}}{1-y_{ij}} \Big[& \hat{\mathcal{R}}_{ij}(\xi_j, y_{ij}) - \hat{\mathcal{R}}_{ij}(0, y_{ij}) \Theta(\xi_c - \xi) \\ & - \hat{\mathcal{R}}_{ij}(\xi_j, 1) \Theta(y - 1 + \delta_O) + \hat{\mathcal{R}}_{ij}(0,1) \Theta(\xi_c - \xi) \Theta(y - 1 + \delta_O) \Big] , \end{aligned} \tag{2.2.20}$$

meaning that the integrals become finite by subtracting the soft limit $\hat{\mathcal{R}}_{ij}(0, y_{ij})$, the collinear limit $\hat{\mathcal{R}}_{ij}(\xi_j, 1)$, and the soft-collinear limit $\hat{\mathcal{R}}_{ij}(0,1)$ from the full divergent squared real amplitude contained in $\hat{\mathcal{R}}_{ij}(\xi_j, y_{ij})$. In the following, we will work out the explicit expressions for the subtraction terms.

Collinear Counterterms

First, we consider the collinear limit. In fig. 2.2 we show an example of a $g \to q\bar{q}$ splitting where we showcase the momentum relations for an FKS-pair $(i,j) \in \mathcal{P}_{\text{FKS}}$. Defining the energy fraction

$$z \equiv \frac{k_j^0}{\bar{k}_i^0} , \quad \text{and thus} \quad 1 - z = \frac{k_i^0}{\bar{k}_i^0} , \tag{2.2.21}$$

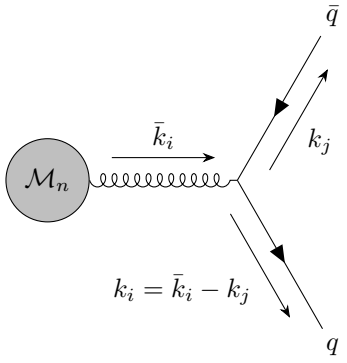


Figure 2.2. Example of a QCD final-state $g \to q\bar{q}$ splitting of a final-state gluon showcasing momentum relations.

and requiring the emitting parton $\mathcal{I}_i$ to be on-shell after the splitting, i.e. $k_i^2 = 0$, we can parametrize the collinearity of $\mathcal{I}_i$ with respect to $\mathcal{I}_j$ by writing[4]

$$k_i^2 = \left(\bar{k}_i^\mu - k_j^\mu\right)^2 = \left[(1-z)\,\frac{k_j^\mu}{z} + k_\perp^\mu + X\eta^\mu\right]^2 = 0\,. \tag{2.2.22}$$

This means that the momentum k_i is parametrized by three terms. First, we have the momentum k_j of the emitted parton, rescaled by the energy fraction k_i^0/k_j^0. We add a component $k_\perp$ transverse to the common direction of the collinear FKS pair (i,j), satisfying the condition $k_\perp k_j = 0$. Furthermore, we add a term X multiplied by an auxiliary light-like vector η, which satisfies the conditions $k_\perp \eta = 0$ and $\eta^2 = 0$. This last term restores on-shellness of the emitted parton.

By applying the mentioned conditions, we can solve eq. (2.2.22) and write the momentum of the emitter before the splitting as [59]

$$\bar{k}_i^\mu = \frac{k_j^\mu}{z} + k_\perp^\mu - \eta^\mu \frac{z k_\perp^2}{2\,(1-z)\,k_j \cdot \eta}\,. \tag{2.2.23}$$

Therefore, in the collinear limit $k_\perp \to 0$, we also have

$$\lim_{k_\perp \to 0} \bar{k}_i^2 \sim \lim_{k_\perp \to 0} \frac{z k_\perp^2}{1-z} = 0\,. \tag{2.2.24}$$

The squared real matrix element in the collinear limit for a splitting of a final-state parton $\bar{\mathcal{I}}_i$ from the underlying Born configuration[5] into a pair of collinear partons $\mathcal{I}_i\mathcal{I}_j$ can then be approximated by [103]

$$\mathcal{R}_{ij}^{\bar{\mathcal{I}}_i \to \mathcal{I}_i\mathcal{I}_j} \simeq \frac{8\pi\alpha_S}{\bar{k}_i^2} \hat{P}^\lambda_{\bar{\mathcal{I}}_i \to \mathcal{I}_i\mathcal{I}_j}(z, k_\perp; \epsilon = 0)\mathcal{B}^{(i)}_\lambda\,. \tag{2.2.25}$$

Here, $\hat{P}^\lambda_{\bar{\mathcal{I}}_i \to \mathcal{I}_i\mathcal{I}_j}$ are the polarized form of the *unregularized Altarelli-Parisi splitting functions* [103] in four dimensions (i.e. $\epsilon = 0$), listed in eqs. (C.1.3) to (C.1.6), where the index λ stands for either two spinor or vector indices. In case of the emitting particle being a gluon, the spin-correlated squared Born matrix element $\mathcal{B}^{(i)}_{\mu\nu}$ is obtained by leaving the spin indices of the emitting gluon uncontracted, i.e. [59]

$$\mathcal{B}^{(i)}_{\mu\nu} = \hat{\mathcal{N}}_B \sum_{\{m\},s_i,s_i'} \mathcal{M}_n(\{m\}, s_i)\mathcal{M}_n^\dagger(\{m\}, s_i')\,(\epsilon_{s_i})^*_\mu \left(\epsilon_{s_i'}\right)_\nu\,, \tag{2.2.26}$$

[4] This is akin to a Sudakov parametrization [102], but follows the slightly different parametrization given in [59] App. B.

[5] We note that the barred notation does not denote an antiquark, even though $\mathcal{I}$ can take the form of the general sign for a quark q or antiquark $\bar{q}$ (thus spawning the ridiculous notation $\bar{q}$ and $\bar{\bar{q}}$, which we avoid.)

with the Born matrix element $\mathcal{M}_n(\{m\}, s_i)$, where we denote as s_i the spin of the emitting gluon and sum over all spins and colors $\{m\}$ of the remaining in- and outgoing legs in the Born configuration, leaving the indices of the emitting gluon uncontracted. The normalization factor $\hat{\mathcal{N}}_B$ takes into account the average over initial-state colors and spins. However, we do not include the symmetry factor for identical particles in the Born final state, since for the real-subtraction terms this actually has to be the symmetry factor for the $(n+1)$-body final state, as given in eq. (2.2.1).

The gluon polarization vectors $(\epsilon_{s_i})_\mu$ follow the HELAS [104] convention and are normalized such that

$$\sum_{\mu,\nu} g^{\mu\nu} \left(\epsilon_{s_i}\right)^*_\mu \left(\epsilon_{s_i'}\right)_\nu = -\delta_{s_i s_i'} \,, \tag{2.2.27}$$

leading to the relation [59]

$$\sum_{\mu,\nu} g^{\mu\nu} \mathcal{B}^{(i)}_{\mu\nu} = -\mathcal{B} \,, \tag{2.2.28}$$

where $\mathcal{B}$ is the squared Born amplitude. Therefore, if $\mathcal{B}$ is known, $\mathcal{B}^{(i)}_{\mu\nu}$ has in general 9 independent matrix elements due to its symmetry. For fermion emitters, the spin structure given by $\hat{P}^{s_i s_i'}_{\bar{\mathcal{I}}_i \to \mathcal{I}_i \mathcal{I}_j}$ is simply $\delta^{s_i s_i'}$, which trivially returns $\mathcal{B}$ when acting on $\mathcal{B}^{(i)}_{s_i s_i'}$.

In summary, we have the following *unregularized* real collinear counterterms for the different final-state splitting possibilities $\bar{\mathcal{I}}_i \to \mathcal{I}_i \mathcal{I}_j$ given as[6] [59]

$$\mathcal{R}^{q\to qg}_{ij}(\xi_j, 1) = \mathcal{R}^{\bar{q}\to \bar{q}g}_{ij}(\xi_j, 1) = \frac{8\pi\alpha_S C_F}{\bar{k}_i^2} \frac{1+(1-z)^2}{z} \mathcal{B}\left(\bar{\Phi}^{ij}_n\right) \,, \tag{2.2.29}$$

$$\mathcal{R}^{g\to q\bar{q}}_{ij}(\xi_j, 1) = \frac{8\pi\alpha_S T_F}{\bar{k}_i^2} \left[-g^{\mu\nu} - 4z(1-z)\frac{k_\perp^\mu k_\perp^\nu}{k_\perp^2}\right] \mathcal{B}^{(i)}_{\mu\nu}\left(\bar{\Phi}^{ij}_n\right) \,, \tag{2.2.30}$$

$$\mathcal{R}^{g\to gg}_{ij}(\xi_j, 1) = \frac{8\pi\alpha_S C_A}{\bar{k}_i^2} \left[-2\left(\frac{z}{1-z} + \frac{1-z}{z}\right) g^{\mu\nu} + 4z(1-z)\frac{k_\perp^\mu k_\perp^\nu}{k_\perp^2}\right] \mathcal{B}^{(i)}_{\mu\nu}\left(\bar{\Phi}^{ij}_n\right) \,, \tag{2.2.31}$$

where the spin correlations $k_\perp^\mu k_\perp^\nu \mathcal{B}_{\mu\nu}$ in eqs. (2.2.30) and (2.2.31) give rise to terms where the squared Born matrix element does not simply factorize. We also note that because of our definition of $k_\perp$ (see eq. (2.2.22)), it is generally possible to write it as $(0, \mathbf{k}_\perp)$, where $\mathbf{k}_\perp$ is perpendicular to the common direction of the FKS pair, and thus any term in the sum over the indices μ, ν with either index refering to the energy component vanishes. The definitions of the Casimir constants C_F and C_A, as well as T_F, will be given in eqs. (2.2.45) to (2.2.47).

[6] Due to a slightly different ansatz in the derivation of eq. (2.2.23), which is inspired by POWHEG [59], the signs in front of the second terms in eqs. (2.2.30) and (2.2.31) differ from the ones found in [103] and in other literature, e.g. [40, 105].

The *regularized* real collinear counterterms $\tilde{\mathcal{R}}_{ij}(\xi_j, 1)$ (see eq. (2.2.4)) simply follow from eqs. (2.2.29) to (2.2.31) and have the form

$$\tilde{\mathcal{R}}_{ij}^{q\to qg}(\xi_j,1) = \tilde{\mathcal{R}}_{ij}^{\bar{q}\to\bar{q}g}(\xi_j,1) = \mathcal{S}_{ij}\frac{16\pi\alpha_S C_F}{(1-z)\,s}\left[1+(1-z)^2\right]\mathcal{B}\left(\bar{\Phi}_n^{ij}\right)\,, \tag{2.2.32}$$

$$\tilde{\mathcal{R}}_{ij}^{g\to q\bar{q}}(\xi_j,1) = \mathcal{S}_{ij}\frac{16\pi\alpha_S T_F}{(1-z)\,s}\left[-zg^{\mu\nu}-4z^2(1-z)\,\frac{k_\perp^\mu k_\perp^\nu}{k_\perp^2}\right]\mathcal{B}_{\mu\nu}^{(i)}\left(\bar{\Phi}_n^{ij}\right)\,, \tag{2.2.33}$$

$$\tilde{\mathcal{R}}_{ij}^{g\to gg}(\xi_j,1) = \mathcal{S}_{ij}\frac{16\pi\alpha_S C_A}{(1-z)\,s}\left[-2\left(\frac{z^2}{1-z}+(1-z)\right)g^{\mu\nu}+4z^2(1-z)\,\frac{k_\perp^\mu k_\perp^\nu}{k_\perp^2}\right]\mathcal{B}_{\mu\nu}^{(i)}\left(\bar{\Phi}_n^{ij}\right)\,, \tag{2.2.34}$$

where we used the relations

$$\bar{k}_i^2 = 2k_i\cdot k_j = 2k_i^0k_j^0\,(1-y_{ij}) = 2z\,(1-z)\,(\bar{k}_i^0)^2\,(1-y_{ij})\,, \tag{2.2.35}$$

and

$$\left(\frac{\xi}{z}\right)^2 = \left(\frac{2k_j^0}{\sqrt{s}}\frac{\bar{k}_i^0}{k_j^0}\right)^2 = \frac{4(\bar{k}_i^0)^2}{s}\,. \tag{2.2.36}$$

The multiplication with $\mathcal{S}_{ij}$ is highly symbolic for eqs. (2.2.32) and (2.2.33), as they are equal to 1 in both cases because of eq. (2.1.29). However, for a $g\to gg$ splitting we have $S_{ij} = h(z_{ij})$, as defined in eq. (2.1.31).

Soft and Soft-Collinear Counterterms

Examining eqs. (2.2.29) and (2.2.31), we can also deduce the soft behavior for the different splittings as either $z\to 0$ or $z\to 1$, corresponding to either emitted or emitting parton becoming soft, respectively. Consequently, only $q\to qg/\bar{q}\to\bar{q}g$ and $g\to gg$ splittings, i.e. the radiation of a soft gluon, show soft singularities.

In the soft limit the terms introducing spin correlations vanish, meaning that the radiation of a soft gluon does not affect the spin of its emitter. More generally, for large enough emission angles the soft gluon cannot resolve the spin structure of the underlying process due to its long wavelength in relation to the rest of the particles in the hard scattering process.[7] While the soft radiation of a gluon also only changes the momentum of the emitter negligibly, it does change the color structure of the system since the gluon carries color charge. Because of the large wavelength of the soft gluon, it can spread the color flow over large distances, which also leads to non-local color correlations [107]. In contrast to the emission of a soft photon in QED, the square of the matrix element therefore does not generally factorize in the soft limit.

However, in the soft-collinear limit, i.e. $y\to 1$ together with either $z\to 0$ or $z\to 1$,

[7] This phenomenon is known in the context of color coherence, see e.g. [106]

the angle between the emitted gluon and its emitter becomes small enough for the gluon to be able to resolve the internal structure of the hard process. Thus, non-local color correlations are not an issue in this case and the squared Born matrix element fully factorizes. For the regularized soft-collinear counterterms $\tilde{\mathcal{R}}_{ij}(0,1)$ from eq. (2.2.20) we then use eqs. (2.2.32) and (2.2.34) since they are both finite in the soft limit where $z \to 0$. Because we have defined $\mathcal{P}_{\text{FKS}}$ in eq. (2.1.11) in such a way that only the emitted parton can become soft, and keeping eq. (2.1.31) in mind, we do not concern ourselves with the limit $z \to 1$ in which the emitter becomes soft.

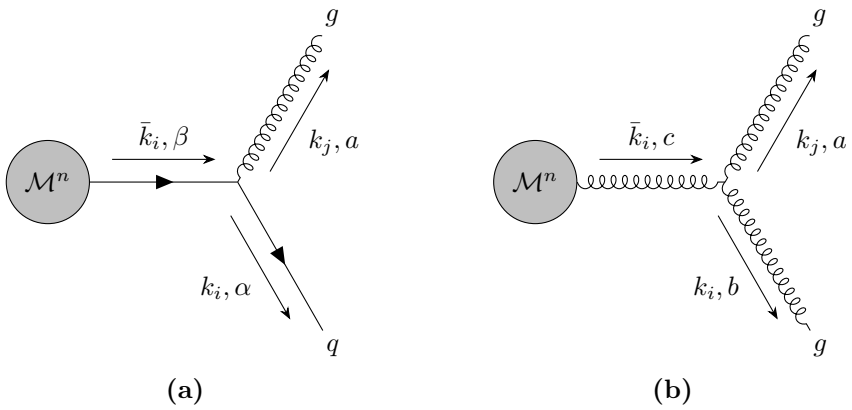


Figure 2.3. Example of QCD FSR of a gluon by an external quark (a) or gluon (b), respectively, with additional greek letters α, β for quark color indices and roman letters a, b, c for the gluon color indices.

Outside of the soft-collinear limit, the computation for the soft limit needs a different ansatz for the reasons stated above. In order to derive the soft counterterm $\tilde{\mathcal{R}}_{ij}(0, y_{ij})$ from eq. (2.2.20), we start by stating the matrix element for the gluon radiation off an external quark with mass m, as depicted in fig. 2.3(a). It is [27]

$$\mathcal{M}^{n+1}_{ij;k}(\Phi_{n+1}) = g_s \hat{T}^a_{\alpha\beta} \epsilon^*_{s_j,\mu}(k_j) \bar{u}(k_i) \gamma^\mu \frac{\not{k}_i + \not{k}_j + m}{(k_i + k_j)^2 - m^2} \mathcal{M}^n(\bar{\Phi}^{ij}_n) \tag{2.2.37}$$

$$= g_s \hat{T}^a_{\alpha\beta} \epsilon^*_{s_j,\mu}(k_j) \bar{u}(k_i) \frac{-(\not{k}_i - m)\gamma^\mu + 2k_i^\mu + \gamma^\mu \not{k}_j}{2k_i \cdot k_j} \mathcal{M}^n(\bar{\Phi}^{ij}_n)\,, \tag{2.2.38}$$

with the appropriate color charge operator $\hat{T}^a_{\alpha\beta}$. Furthermore, we denote $\not{k} = \gamma^\mu k_\mu$ in the usual slash notation and used the anticommutator relation $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$ to write the second line. Additionally, we use an index k for the $(n+1)$-body matrix element for reasons that will become obvious in eq. (2.2.40). The first term in the numerator then gives the Dirac equation $\bar{u}(k_i)(\not{k}_i - m) = 0$ and thus vanishes. In the soft gluon limit, i.e. $k_j^0 \to 0$, we write $k_j = k_j^0 \hat{k}_j$ (see eq. (2.1.39)) and disregard any terms only proportional to k_j^0 while keeping those proportional to k_i^0. In addition, since

the radiation of a soft gluon does not change the momentum of the emitter, we can express k_i in the underlying Born kinematics as $\bar{k}_i$. Neglecting recoil effects in this manner is commonly known as the *eikonal approximation* (see e.g. [106, 108]) in which we can write the matrix element for the radiation of a soft gluon as

$$\mathcal{M}^{n+1}_{ij;k}(\Phi_{n+1}) \simeq g_s \hat{T}^a_{\alpha\beta} \epsilon^*_{s_j,\mu}(k_j) \frac{1}{k_j^0} \frac{\bar{k}_i^\mu}{\bar{k}_i \cdot \hat{k}_j} \mathcal{M}^n(\bar{\Phi}^{ij}_n)\,, \tag{2.2.39}$$

where we absorbed the spinor $\bar{u}(\bar{k}_i)$ into the matrix element. The singular factor $1/k_j^0$ will cancel against a factor ξ in the regularized squared amplitude $\tilde{R}_{ij}$, which we address later. Because the gluon cannot resolve the spin structure of the underlying process due to its relatively long wavelength, the form of the eikonal approximation is universal, i.e. the same for a quark, antiquark, or gluon emitter.

As stated earlier, another effect of the large wavelength of the soft gluon is the fact that it can spread the color flow over large distances, which also leads to non-local color correlations. We thus have to sum over all momenta and color operators of the external hard partons such that[8]

$$\mathcal{M}^{n+1}_{ij}(\Phi_{n+1}) \simeq g_s \epsilon^*_{s_j,\mu}(k_j) \frac{1}{k_j^0} \left(\sum_{k=1}^{n+2} \hat{T}^a_{c_k,c'_k} \frac{\bar{k}_k^\mu}{\bar{k}_k \cdot \hat{k}_j} \right) \mathcal{M}^n(\bar{\Phi}^{ij}_n)\,. \tag{2.2.40}$$

Here, we introduced the general notation $\hat{T}^a_{c_k,c'_k}$ for the color charge operators, with their action on color space for an emitted gluon with color a dependent on the nature of its emitter $\mathcal{I}_k$. Decomposing the Born matrix element in color space gives

$$\mathcal{M}^n(\bar{\Phi}^{ij}_n) = \left\langle \mathcal{M}^n(\bar{\Phi}^{ij}_n) \middle| c_1, \cdots, c_n \right\rangle\,, \tag{2.2.41}$$

with the abstract color space basis $\{\,|c_1, \cdots, c_n\rangle\}$. For a final-state emitter the exact actions on color space are then given by

$$\hat{T}^a_{\alpha\beta}\ |c_1, \cdots, \alpha_q, \cdots, c_n\rangle = \ \ +t^a_{\alpha\beta}\ |c_1, \cdots, \beta_q, \cdots, c_n, a\rangle\,, \tag{2.2.42}$$

$$\hat{T}^a_{\alpha\beta}\ |c_1, \cdots, \alpha_{\bar{q}}, \cdots, c_n\rangle = \ \ -t^a_{\beta\alpha}\ |c_1, \cdots, \beta_{\bar{q}}, \cdots, c_n, a\rangle\,, \tag{2.2.43}$$

$$\hat{T}^a_{bc}\ |c_1, \cdots, b_g, \cdots, c_n\rangle = -if^{abc}\ |c_1, \cdots, c_g, \cdots; c_n, a\rangle\,. \tag{2.2.44}$$

With $t^a_{\alpha\beta}$ we denote the color matrices in the fundamental representation of $SU(3)$. For an initial-state quark or antiquark the action would be exactly opposite to eqs. (2.2.42) and (2.2.43). The operator $\hat{T}^a_{bc}$ is in the adjoint representation of $SU(3)$, with the corresponding structure constants f^{abc}.

[8] In fact, we can just sum over all external legs, since the particular term for an external particle that is not strongly interacting is just equal to zero.

Generally, in $SU(N)$ we have the relations

$$\mathrm{tr}\left[t^a t^b\right] = T_F \delta^{ab}\,, \quad \frac{1}{2} = T_F \tag{2.2.45}$$

$$\frac{1}{N}\mathrm{tr}\sum_{a,b} t^a t^b = \frac{1}{N}\delta^{ab} T_F \delta_{ab} = \frac{N^2-1}{2N} = C_F \tag{2.2.46}$$

$$\sum_{c,d} f^{acd} f^{bcd} = C_A \delta^{ab}\,, \quad N = C_A\,. \tag{2.2.47}$$

We now simply square eq. (2.2.40) and sum over all final-state spins and colors. The regularized squared amplitude from eq. (2.2.4) in the soft limit $\xi_j = 0$ is then [42, 59]

$$\tilde{\mathcal{R}}_{ij}(0, y_{ij}) = \delta_{g\mathcal{I}_j}\,(1 - y_{ij})\,\mathcal{S}_{ij}^{\mathrm{soft}} \frac{16\pi\alpha_S}{s} \sum_{k,l=1}^{n+2} \frac{\bar{k}_k \cdot \bar{k}_l}{(\bar{k}_k \cdot \hat{k}_j)(\bar{k}_l \cdot \hat{k}_j)} \mathcal{B}_{kl}(\bar{\Phi}_n^{ij})\,, \quad (i,j) \in \mathcal{P}_{\mathrm{FKS}}\,. \tag{2.2.48}$$

We note that we multiply with the soft $\mathcal{S}$-functions, where we use the definitions of the d_{ij} in the soft limit from section 2.1.2, and that we used $\delta_{g\mathcal{I}_j}$ to make explicit that only gluon radiation needs a soft subtraction term. Also, we used the fact that $\xi_j = 2k_j^0/\sqrt{s}$ and thus

$$\left(\frac{\xi}{k_j^0}\right)^2 = \frac{4}{s} \tag{2.2.49}$$

completely cancels the $1/(k_j^0)^2$ singularity in the limit $k_j^0 \to 0$. Furthermore, we recall the gluon polarization vector normalization as stated in eq. (2.2.27). In Feynman gauge the gluon polarization sum then is

$$\sum_{s_j} \epsilon_{s_j,\mu}(k_j)\epsilon^*_{s_j,\nu}(k_j) = -g_{\mu\nu}\,. \tag{2.2.50}$$

The resulting minus sign for the squared amplitude is then absorbed into the definition of the color-correlated squared Born amplitude $\mathcal{B}_{kl}$, which is given as

$$\begin{aligned}\mathcal{B}_{kl}(\bar{\Phi}_n^{ij}) = -\hat{\mathcal{N}}_B \sum_{\substack{\text{spins}\\ \text{colors}}} &\left\langle \mathcal{M}^n(\bar{\Phi}_n^{ij}) \middle| c_1, \cdots, c_k, \cdots, c_l, \cdots, c_n, a \right\rangle \\ &\left\langle c_1, \cdots, c'_k, \cdots, c'_l, \cdots, c_n, a \middle| \mathcal{M}^n(\bar{\Phi}_n^{ij}) \right\rangle T^a_{c_k,c'_k} T^a_{c_l,c'_l}\,,\end{aligned} \tag{2.2.51}$$

where we summarized the color matrices and structure constants from eqs. (2.2.42) to (2.2.44) symbolically as $T^a_{c,c'}$. We sum over all repeating color indices and remaining spins. The normalization factor $\hat{\mathcal{N}}_B$ takes into account the average over initial-state color and spins. We note that we again do not include the symmetry factor for identical

particles in the Born final state, since for the real-subtraction terms this actually has to be the symmetry factor for the $(n+1)$-body final state, as given in eq. (2.2.1). Furthermore, color conservation and the relations in eqs. (2.2.46) and (2.2.47) for the Casimir constants $C(\bar{\mathcal{I}}_l)$ for the emitting flavor $\bar{\mathcal{I}}_l$ gives [59]

$$\sum_{k,k\neq l}^{n+2} \mathcal{B}_{kl} = C(\bar{\mathcal{I}}_l)\mathcal{B}\,. \tag{2.2.52}$$

With this relation, together with the assumption that $\mathcal{B}$ is known and the fact that $\mathcal{B}_{kl}$ is symmetric – which leaves us with

$$n_{\text{indep.}} = \frac{n_S\,(n_S-1)}{2} \tag{2.2.53}$$

independent squared matrix elements – we can rewrite eq. (2.2.48) such that [59]

$$\begin{aligned}\tilde{\mathcal{R}}_{ij}(0,y_{ij}) = \delta_{g\mathcal{I}_j}\,(1-y_{ij})\,\mathcal{S}_{ij}^{\text{soft}}\frac{16\pi\alpha_S}{s}\Bigg[&\sum_{k>l}^{n+2}\frac{2\bar{k}_k\cdot\bar{k}_l}{(\bar{k}_k\cdot\hat{k}_j)(\bar{k}_l\cdot\hat{k}_j)}\mathcal{B}_{kl}(\bar{\Phi}_n^{ij}) \\ &-\mathcal{B}(\bar{\Phi}_n^{ij})\sum_{k=1}^{n+2}\frac{\bar{k}_k^2}{(\bar{k}_k\cdot\hat{k}_j)^2}C(\bar{\mathcal{I}}_k)\Bigg]\,,\quad (i,j)\in\mathcal{P}_{\text{FKS}}\,.\end{aligned} \tag{2.2.54}$$

2.2.2. Hadron Collisions and Initial-state Subtraction Terms

So far we have only looked at FSR in the partonic process. Of course, for leptonic initial states this is perfectly sufficient when considering NLO QCD corrections. A correct physical description for hadron collisions, however, is generally only possible by also taking into account the effects of PDFs and ISR. In general, the factorization theorem [109–111] states that for the collision of two hadrons $H_\oplus$ with momenta $K_\oplus$ the hadronic differential cross section $d\sigma_H$ factorizes into a hard cross section $d\sigma$, containing all the small distance, high energy interactions, and the parametrization of the low energy contribution below an arbitrary scale μ_F given by the PDF $f_{\mathcal{I}_\oplus}^{H_\oplus}$ such that

$$d\sigma_H(K_\oplus,K_\ominus) = \sum_{a,b}\int dx_\oplus dx_\ominus f_a^{H_\oplus}(x_\oplus,\mu_F) f_b^{H_\ominus}(x_\ominus,\mu_F) d\sigma_{ab}(k_\oplus,k_\ominus)\,, \tag{2.2.55}$$

where we introduced the shorthand notation for the parton flavors

$$a\equiv\mathcal{I}_\oplus \quad\text{and}\quad \bar{a}\equiv\bar{\mathcal{I}}_\oplus\,,\quad b\equiv\mathcal{I}_\ominus \quad\text{and}\quad \bar{b}\equiv\bar{\mathcal{I}}_\ominus\,, \tag{2.2.56}$$

for the sake of readability.[9] The notation $d\sigma_{ab}$ then denotes the differential cross section for hard processes with an initial-state ab, where we imply a sum over all possible final states, if not stated otherwise. Furthermore, the momenta $k_\oplus$ of the partons $\mathcal{I}_\oplus$ entering the hard interaction are related to the hadronic momenta by

$$k_\oplus = x_\oplus K_\oplus\,, \quad \text{with} \quad x_\oplus K_\oplus + x_\ominus K_\ominus = \sum_{l=3}^{n+3} k_l\,, \tag{2.2.57}$$

where $x_\oplus$ is the momentum fraction of the incoming partons.

At NLO, however, without modifications to the PDFs the real and virtual IR divergences do not cancel completely and we are left with initial-state collinear divergences. Thus, renormalized PDFs are introduced at the next order of α_S which take the form [112, 113]

$$\tilde{f}_{\bar{a}}^{H_\oplus}(\bar{x}_\oplus, \mu_F) \equiv \int_0^1 \int_0^1 dx_\oplus dz f_a^{H_\oplus}(x_\oplus)\Gamma_{a\leftarrow\bar{a}}(z,\mu_F)\delta(\bar{x}_\oplus - z x_\oplus) \tag{2.2.58}$$

$$= \int_{\bar{x}_\oplus}^1 \frac{dz}{z} f_a^{H_\oplus}\left(\frac{\bar{x}_\oplus}{z}\right)\Gamma_{a\leftarrow\bar{a}}(z,\mu_F)\,. \tag{2.2.59}$$

with the $\ominus$-direction being completely analogous. Here, $\bar{x}_\oplus$ is the momentum fraction that remains for the parton $\bar{a}$ entering the hard interaction after a parton a splits into $\bar{a}$ and an additional emitted parton $\mathcal{I}_j$, where $x_\oplus$ is the momentum fraction of parton a before radiation. This is illustrated in fig. 2.4. The relation between the fractions is then given by [67] (see appendix A.2 for more details)

$$\bar{x}_\oplus = z_\oplus x_\oplus\,, \quad z_\oplus \equiv \sqrt{1-\xi}\sqrt{\frac{2-\xi\,(1\pm y)}{2-\xi\,(1\mp y)}}\,, \tag{2.2.60}$$

where we always drop the index $\oplus$ of z, as it can be understood by context. For either case of an initial-state collinear singularity, $y = 1$ and $y = -1$, we then have

$$\lim_{y\to\pm 1} z = 1 - \xi\,. \tag{2.2.61}$$

Furthermore, we note that the PDF $f_{\bar{a}}^{H_\oplus}$ in eq. (2.2.59) depends on the nature of parton $\bar{a}$ and can therefore have a different flavor content than the unrenormalized PDF $f_a^{H_\oplus}$.

Coming back to eq. (2.2.59), the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution kernel $\Gamma_{a\leftarrow\bar{a}}$, as well as the equation itself, is to be understood as a backwards evolution of the PDF from the energy scale $\bar{x}$ up to that of the hadron, taking into

[9] Again, the barred notation here does not denote an antiparticle.

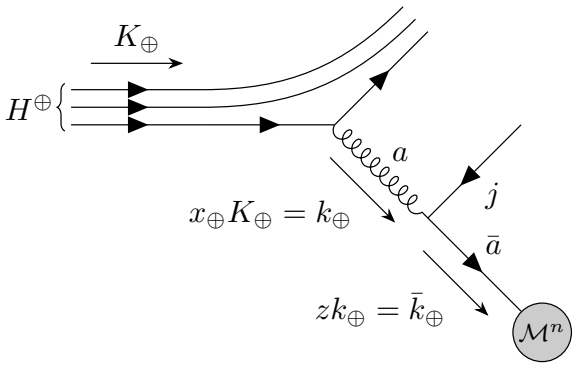


Figure 2.4. Example for a hadron $H^\oplus$ with momentum $K_\oplus$ in $\oplus$-direction radiating a gluon with momentum $k_\oplus$, which then splits into a quark-antiquark pair, of which the quark enters the hard interaction with momentum $\bar{k}_\oplus$.

account the radiation at scale x. At NLO, the kernel is given by [112, 113]

$$\Gamma_{a\leftarrow\bar{a}}(z,\mu_F) \equiv \delta_{a\bar{a}}\delta(1-z) - \frac{\alpha_S}{2\pi}\left(\frac{1}{\bar{\epsilon}}P_{\bar{a}\to a\mathcal{I}_j}(z,0) - K_{\bar{a}\to a\mathcal{I}_j}(z)\right) + \mathcal{O}(\alpha_S^2)\,, \quad (2.2.62)$$

with the *regularized* Altarelli-Parisi kernels $P_{\bar{a}\to a\mathcal{I}_j}(z,0)$ in four dimensions for a splitting of parton $\bar{a}$ into parton a and an emitted parton $\mathcal{I}_j$, viewed as an evolution backwards in time. The exact definition of $P_{\bar{a}\to a\mathcal{I}_j}(z,0)$ will be given in eq. (2.3.43) in the context of section 2.3.3. We note that in the $\overline{\text{MS}}$ renormalization scheme the term $K_{\bar{a}\to a\mathcal{I}_j}$ is exactly zero and thus we will neglect it in the following. Furthermore, we introduced the $\overline{\text{MS}}$-scheme pole

$$\frac{1}{\bar{\epsilon}} \equiv \frac{1}{\epsilon}\frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}\left(\frac{\mu_R^2}{\mu_F^2}\right)^\epsilon + \mathcal{O}(\epsilon)\,, \quad (2.2.63)$$

where the gamma function $\Gamma(1-\epsilon)$ should not be confused with the evolution kernel $\Gamma_{a\leftarrow\bar{a}}$. In the scope of this work, we will not consider cases where the renormalization scale μ_R and factorization scale μ_F differ, thus

$$\mu = \mu_R = \mu_F\,.$$

The second term in eq. (2.2.62) acts as a subtraction term for the remnant collinear divergences due to the bare PDF in eq. (2.2.55) at NLO and is collinear divergent itself. In fact, due to the universality of eqs. (2.2.55) and (2.2.62), we can interpret the evolution kernel $\Gamma_{a\leftarrow\bar{a}}$ as a parton density. We then call the partonic differential cross section $d\sigma_{ab}$ from the right-hand side of eq. (2.2.55) the *unsubtracted* differential cross section, and

$$d\hat{\sigma}_{ab} \equiv \Gamma_{a\leftarrow\bar{a}}\,\Gamma_{b\leftarrow\bar{b}}\,d\sigma_{cd} \quad (2.2.64)$$

the *subtracted* partonic differential cross section. With the perturbative expansion at

NLO written out for both for LO (0) and NLO (1) as

$$d\sigma_{ab} = d\sigma_{ab}^{(0)} + d\sigma_{ab}^{(1)}\,, \quad d\hat{\sigma}_{ab} = d\hat{\sigma}_{ab}^{(0)} + d\hat{\sigma}_{ab}^{(1)}\,, \tag{2.2.65}$$

the relation between them at NLO then is [41]

$$d\hat{\sigma}_{ab}^{(0)}(\bar{k}_\oplus, \bar{k}_\ominus) = d\sigma_{ab}^{(0)}(\bar{k}_\oplus, \bar{k}_\ominus) \tag{2.2.66}$$

$$d\hat{\sigma}_{ab}^{(1)}(k_\oplus, k_\ominus) = d\sigma_{ab}^{(1)}(k_\oplus, k_\ominus) + \underbrace{\frac{\alpha_S}{2\pi}\sum_{\bar{a}}\int dz \frac{1}{\bar{\epsilon}} P_{\bar{a}\to a\mathcal{I}_j}(z,0) d\sigma_{\bar{a}b}^{(0)}(zk_\oplus, k_\ominus)}_{d\sigma_{ab}^{(\mathrm{cnt},\oplus)}}$$
$$+ \underbrace{\frac{\alpha_S}{2\pi}\sum_{\bar{b}}\int dz \frac{1}{\bar{\epsilon}} P_{\bar{b}\to b\mathcal{I}_j}(z,0) d\sigma_{a\bar{b}}^{(0)}(k_\oplus, zk_\ominus)}_{d\sigma_{ab}^{(\mathrm{cnt},\ominus)}}\,, \tag{2.2.67}$$

where we sum over all possible splittings leading to a real initial-state flavor structure given by $d\hat{\sigma}_{ab}^{(1)}$, and consequently all corresponding underlying Born flavor combinations. The PDF counterterms $d\sigma_{ab}^{(\mathrm{cnt},\oplus)}$ (for massless initial-state particles) have collinear singularities and exactly cancel those left by the partial cancellation of divergences between the real and virtual contributions mentioned earlier.

In terms of numerical calculations and FKS, we then need adequate terms to treat the divergences in both the $d\sigma_{ab}^{(\mathrm{cnt},\oplus)}$ and, analogous to the FSR case, the partonic cross section $d\sigma_{ab}^{(1)}$. Fortunately, the needed terms naturally emerge when analogously using the prescriptions described in section 2.2.1. For this, we will rewrite eq. (2.2.4) to account for both singularities at $y = \pm 1$ such that [67]

$$\mathcal{R}_\alpha = \frac{1}{\xi_j^2}\frac{1}{1-y_j^2}\tilde{\mathcal{R}}_{ij}(\xi_j, y_j)\,, \quad \tilde{\mathcal{R}}_{ij}(\xi_j, y_j) \equiv \left[\xi_j^2\left(1-y_j^2\right)\mathcal{R}_\alpha\right]\,. \tag{2.2.68}$$

As stated before, the singularities in the PDF counterterms in eq. (2.2.67) are purely collinear in nature. Therefore, the expansion in ξ introduced in eq. (2.2.10) still applies, while we neglect the first term of the expansion as it does not give new contributions. With the definition of eq. (2.2.68), the integrand of eq. (2.2.5) becomes

$$\mathcal{R}_\alpha^{(\mathrm{in})}(f_r) \equiv \mathcal{P}_\mathrm{c}(\xi_j)\left(1-y_j^2\right)^{-1-\epsilon}\tilde{\mathcal{R}}_{ij}(\xi_j, y_j)\,, \tag{2.2.69}$$

with

$$\mathcal{P}_\mathrm{c}(\xi_j) \equiv \left(\frac{1}{\xi_j}\right)_\mathrm{c} - 2\epsilon\left(\frac{\log\xi_j}{\xi_j}\right)_\mathrm{c}\,. \tag{2.2.70}$$

Using eq. (2.2.11), we regularize the remaining collinear divergences by writing [41]

$$\begin{aligned}\left(1-y_j^2\right)^{-1-\epsilon} =& -\frac{(2\delta_I)^{-\epsilon}}{2\epsilon}\left[\delta(1-y_j)+\delta(1+y_j)\right] \\ &+\frac{1}{2}\left[\left(\frac{1}{1-y_j}\right)_{\delta_I}+\left(\frac{1}{1+y_j}\right)_{\delta_I}\right]+\mathcal{O}(\epsilon)\,,\end{aligned} \tag{2.2.71}$$

and can thus represent eq. (2.2.69) as the sum

$$\mathcal{R}_\alpha^{(\text{in})}(f_r) = \mathcal{R}_\alpha^{(\text{in},\oplus)} + \mathcal{R}_\alpha^{(\text{in},\ominus)} + \mathcal{R}_\alpha^{(\text{in,fin})} + \mathcal{O}(\epsilon) \tag{2.2.72}$$

where

$$\mathcal{R}_\alpha^{(\text{in,fin})}(f_r) \equiv \mathcal{P}_\text{c}(\xi_j)\frac{1}{2}\left[\left(\frac{1}{1-y_j}\right)_{\delta_I}+\left(\frac{1}{1+y_j}\right)_{\delta_I}\right]\tilde{\mathcal{R}}_{ij}(\xi_j,y_j)\,, \tag{2.2.73}$$

$$\mathcal{R}_\alpha^{(\text{in},\oplus)}(f_r) \equiv -\frac{(2\delta_I)^{-\epsilon}}{2\epsilon}\delta\left(1\mp y_j\right)\mathcal{P}_\text{c}(\xi_j)\tilde{\mathcal{R}}_{ij}(\xi_j,y_j)\,. \tag{2.2.74}$$

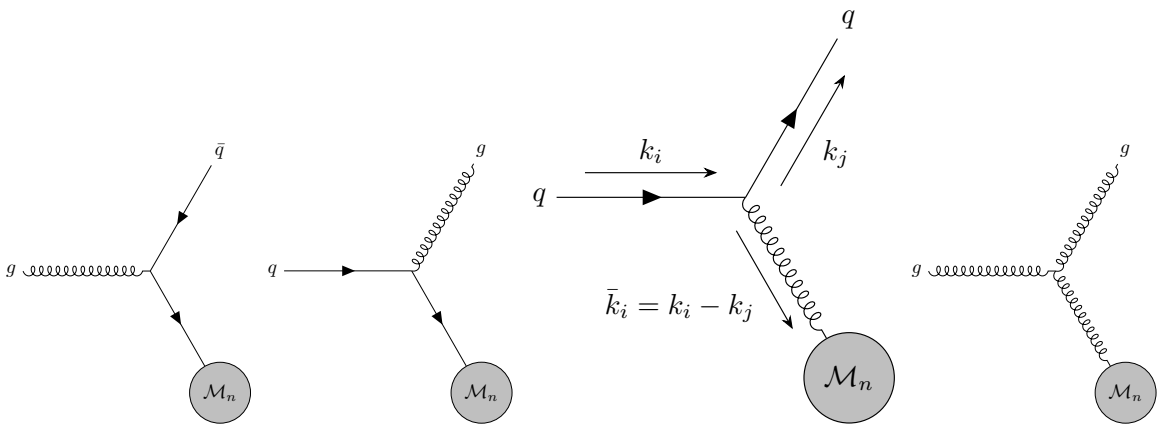
The term $\mathcal{R}_\alpha^{(\text{in},\oplus)}$ only exists in the collinear limit and contains an ϵ-pole that can be associated with a soft singularity. It cancels corresponding terms in the bare virtual contribution, and its counterterms after integrating out the radiation phase space. We will thus deal with it in the context of section 2.3.

In constrast, the term $\mathcal{R}_\alpha^{(\text{in,fin})}$ is finite and the analogon to eq. (2.2.20) for ISR. Here, we get additional collinear and soft-collinear terms for $y_j = -1$, however, which is unique to the initial-state case. In the following, we give an overview of all the changes in the subtraction terms compared to the FSR case.

Collinear Counterterms

For the initial-state collinear counterterms we have some differences to the final-state case due to the different momentum structure and additional splittings. The possible splittings and an example for the momenta are given in fig. 2.5.

Any momentum configuration in which a massless final-state parton becomes collinear to one of the massless incoming partons can be expressed by the underlying Born momenta plus the additional degree of freedom in the energy of the radiated collinear



(a) $g\bar{q} \leftarrow q$ **(b)** $qg \leftarrow q$ **(c)** $qq \leftarrow g$ **(d)** $gg \leftarrow g$

Figure 2.5. All possible QCD initial-state splittings $\mathcal{I}_i\mathcal{I}_j \leftarrow \bar{\mathcal{I}}_i$ of an emitter $\bar{\mathcal{I}}_i$ from the underlying Born configuration into an FKS pair $\mathcal{I}_i\mathcal{I}_j$ are shown. The direction of the arrow in the splitting notation stresses the fact that the splitting for ISR is seen as a backwards evolution.

parton $\mathcal{I}_j$, parametrized by z, as [67]

$$\mathbf{\Phi}^{(\alpha)}_{n,\oplus} \equiv \left\{x_\oplus, x_\ominus, z, \bar{k}_3, \cdots, \bar{k}_{n+2}\right\}, \quad z x_\oplus K_\oplus + x_\ominus K_\ominus = \sum_{l=3}^{n+2} \bar{k}_l\,, \tag{2.2.75}$$

$$\mathbf{\Phi}^{(\alpha)}_{n,\ominus} \equiv \left\{x_\oplus, x_\ominus, z, \bar{k}_3, \cdots, \bar{k}_{n+2}\right\}, \quad x_\oplus K_\oplus + z x_\ominus K_\ominus = \sum_{l=3}^{n+2} \bar{k}_l\,. \tag{2.2.76}$$

Therefore, this parametrization effectively represents an $(n+1)$-body kinematics with reduced degrees of freedom for the radiated parton. The underlying Born configuration then is

$$\bar{\mathbf{\Phi}}^{(\alpha)}_{n} \equiv \left\{\bar{x}_\oplus, \bar{x}_\ominus, \bar{k}_3, \cdots, \bar{k}_{n+2}\right\}, \tag{2.2.77}$$

where $\bar{x}_\oplus$ is defined as in eq. (2.2.60). The phase-space element for the initial-state collinear radiation is then given by

$$d\mathbf{\Phi}^{(\alpha)}_{n,\oplus} \equiv dx_\oplus\, dx_\ominus\, dz\, d\bar{\Phi}^{(\alpha)}_n \left(z k_\oplus + k_\ominus; \bar{k}_3, \cdots, \bar{k}_{n+2}\right), \tag{2.2.78}$$

$$d\mathbf{\Phi}^{(\alpha)}_{n,\ominus} \equiv dx_\oplus\, dx_\ominus\, dz\, d\bar{\Phi}^{(\alpha)}_n \left(k_\oplus + z k_\ominus; \bar{k}_3, \cdots, \bar{k}_{n+2}\right). \tag{2.2.79}$$

Hence, the relation between the degenerate $(n+1)$-body phase space and the underlying Born phase space is

$$d\mathbf{\Phi}^{(\alpha)}_{n,\oplus} = d\bar{\mathbf{\Phi}}^{(\alpha)}_{n} \frac{dz}{z}\,, \tag{2.2.80}$$

where the Jacobian for the transformation $\bar{x}_\oplus \to x_\oplus$ gives the factor $1/z$.

Additionally, we note that the flavor content of the PDF for the incoming radiating leg is that of the parton before emission, as given in eq. (2.2.59). As stated there as well, in the pure collinear limit (i.e. $z < 1$) this PDF is rescaled in terms of the momentum fraction.

Similar to the derivation of eq. (2.2.23), we can derive the momentum $\bar{k}_i$ of the initial-state parton $\bar{\mathcal{I}}_i$ that enters the hard process after the radiation of an additional parton. Here we have to keep in mind, however, that the virtuality of $\bar{\mathcal{I}}_i$ is negative, i.e. $\bar{k}_i^2 < 0$. With fig. 2.5 as a reference, we thus start from

$$\bar{k}_i = k_i^\mu - k_j^\mu = z k_i^\mu + k_\perp^\mu + X\eta^\mu \,, \tag{2.2.81}$$

where we used the analogous definitions of $k_\perp$ and η to those in eq. (2.2.23) such that $k_\perp k_i = 0$, $k_\perp \eta = 0$, as well as $\eta^2 = 0$. Solving $k_j^2 = 0$ for X and inserting it in eq. (2.2.81) we then obtain

$$\bar{k}_i^\mu = z k_i^\mu + k_\perp^\mu + \eta^\mu \frac{k_\perp^2}{2\,(1-z)\,k_i \cdot \eta} \,. \tag{2.2.82}$$

The resulting *unregularized* real collinear counterterms for the different initial-state splitting possibilities $\mathcal{I}_i\mathcal{I}_j \leftarrow \bar{\mathcal{I}}_i$, as defined in fig. 2.5, are then given by [59]

$$\mathcal{R}_{ij}^{g\bar{q}\leftarrow q}(\xi_j, 1) = \mathcal{R}_{ij}^{gq\leftarrow \bar{q}}(\xi_j, 1) = \frac{8\pi\alpha_S T_F}{-\bar{k}_i^2}\left(z^2 + (1-z)^2\right)\mathcal{B}\left(\bar{\mathbf{\Phi}}_n^{(\alpha)}\right) \,, \tag{2.2.83}$$

$$\mathcal{R}_{ij}^{qg\leftarrow q}(\xi_j, 1) = \mathcal{R}_{ij}^{\bar{q}g\leftarrow \bar{q}}(\xi_j, 1) = \frac{8\pi\alpha_S C_F}{-\bar{k}_i^2}\frac{(1+z^2)}{1-z}\mathcal{B}\left(\bar{\mathbf{\Phi}}_n^{(\alpha)}\right) \,, \tag{2.2.84}$$

$$\mathcal{R}_{ij}^{qq\leftarrow g}(\xi_j, 1) = \frac{8\pi\alpha_S C_F}{-\bar{k}_i^2}\left[-z g^{\mu\nu} + \frac{4\,(1-z)}{z}\frac{k_\perp^\mu k_\perp^\nu}{k_\perp^2}\right]\mathcal{B}_{\mu\nu}^{(i)}\left(\bar{\mathbf{\Phi}}_n^{(\alpha)}\right) \,, \tag{2.2.85}$$

$$\mathcal{R}_{ij}^{gg\leftarrow g}(\xi_j, 1) = \frac{8\pi\alpha_S C_A}{-\bar{k}_i^2}\left[-2\left(\frac{z}{1-z} + z\,(1-z)\right)g^{\mu\nu} + \frac{4\,(1-z)}{z}\frac{k_\perp^\mu k_\perp^\nu}{k_\perp^2}\right]\mathcal{B}_{\mu\nu}^{(i)}\left(\bar{\mathbf{\Phi}}_n^{(\alpha)}\right) \,, \tag{2.2.86}$$

where the minus signs in the denominator of the prefactors have to be included due to the aforementioned fact that in the initial-state case the virtuality $\bar{k}_i^2$ of the emitter is negative. The *regularized* real collinear counterterms $\tilde{\mathcal{R}}_{ij}(\xi_j, 1)$ (see eq. (2.2.4)) follow from eqs. (2.2.84) to (2.2.86) and have the form

$$\tilde{\mathcal{R}}_{ij}^{g\bar{q}\leftarrow q}(\xi_j, 1) = \tilde{\mathcal{R}}_{ij}^{gq\leftarrow \bar{q}}(\xi_j, 1) = \frac{8\pi\alpha_S T_F}{\left(k_i^0\right)^2}(1-z)\left(z^2 + (1-z)^2\right)\mathcal{B}\left(\bar{\mathbf{\Phi}}_n^{(\alpha)}\right) \,, \tag{2.2.87}$$

$$\tilde{\mathcal{R}}_{ij}^{qg\leftarrow q}(\xi_j, 1) = \tilde{\mathcal{R}}_{ij}^{\bar{q}g\leftarrow \bar{q}}(\xi_j, 1) = \frac{8\pi\alpha_S C_F}{\left(k_i^0\right)^2}\left(1+z^2\right)\mathcal{B}\left(\bar{\mathbf{\Phi}}_n^{(\alpha)}\right) \,, \tag{2.2.88}$$

$$\tilde{\mathcal{R}}_{ij}^{qq\leftarrow g}(\xi_j,1) = \frac{8\pi\alpha_S C_F}{\left(k_i^0\right)^2}\left[-z\,(1-z)\,g^{\mu\nu} + 4\frac{(1-z)^2}{z}\frac{k_\perp^\mu k_\perp^\nu}{k_\perp^2}\right]\mathcal{B}_{\mu\nu}^{(i)}\left(\bar{\mathbf{\Phi}}_n^{(\alpha)}\right)\,, \tag{2.2.89}$$

$$\tilde{\mathcal{R}}_{ij}^{gg\leftarrow g}(\xi_j,1) = \frac{8\pi\alpha_S C_A}{\left(k_i^0\right)^2}\left[-2\left(z+z\,(1-z)^2\right)g^{\mu\nu} + 4\frac{(1-z)^2}{z}\frac{k_\perp^\mu k_\perp^\nu}{k_\perp^2}\right]\mathcal{B}_{\mu\nu}^{(i)}\left(\bar{\mathbf{\Phi}}_n^{(\alpha)}\right)\,, \tag{2.2.90}$$

where we used the relations $\lim_{\mathbf{k}_i\parallel\mathbf{k}_j}\xi_j = 1-z$ and

$$\bar{k}_i^2 = (k_i-k_j)^2 = -2k_i\cdot k_j = -2k_i^0k_j^0\,(1\mp y_j) = -2\left(k_i^0\right)^2(1-z)\,(1\mp y_j)\,, \tag{2.2.91}$$

cancelling one order of ξ_j, as well as a factor $(1\mp y_j)$ of $\left(1-y_j^2\right)$ in the regularized squared amplitude, with the remaining one simply becoming a factor of 2. Also, we do not have to multiply $\tilde{\mathcal{R}}_{ij}(\xi_j,1)$ with the $\mathcal{S}$-functions, since they are always equal to one in the initial-state collinear case, as given by eqs. (2.1.28) and (2.1.29).

Soft and Soft-Collinear Counterterms

Looking at the unregularized collinear counterterms in eqs. (2.2.83) to (2.2.86), it becomes obvious that in the initial-state case only $qg \leftarrow q$ (or $\bar{q}g \leftarrow \bar{q}$) and $gg \leftarrow g$ splittings can induce a soft singularity, while the remaining splittings either vanish or remain finite for $\xi \to 0$. Because of the universality of the eikonal approximation in the purely soft limit, the subtraction terms given in eq. (2.2.54) hold true for ISR, after replacing

$$(1-y_{ij}) \to \left(1-y_j^2\right)\,, \tag{2.2.92}$$

and using the initial-state d_{ij}^{soft} from eqs. (2.1.44) to (2.1.46) in the calculation of the $\mathcal{S}$-functions. For the same reasons as stated for the final-state soft-collinear counterterms in section 2.2.1, in the initial-state soft-collinear case we can use the regularized eqs. (2.2.87) to (2.2.90) with $z=1$ (i.e. $\xi_j = 0$).

Finally, the radiation of a soft or soft-collinear parton does neither change the flavor of the incoming parton nor its momentum fraction. Therefore, for the splittings in figs. 2.5(b) and 2.5(d) we have

$$\bar{x}_\oplus = x_\oplus\,, \tag{2.2.93}$$

and the flavor content of the PDF for the radiating leg does not change for the emission of a gluon.

2.3. Integrated Counterterms and the Virtual Cross Section

We introduced the counterterms for the $(n+1)$-body kinematics in section 2.2. What remains are the details of the computation of the integrated counterterms, which are given in eq. (2.2.17) and which regularize the virtual contribution as in eq. (2.0.3).

According to the KLN theorem, in dimensional regularization all ϵ-poles due to IR divergences cancel in the sum of all contributions to the total cross section. Therefore, assuming all UV divergences have been renormalized, the IR divergent part of the squared *bare* one-loop amplitude $\mathcal{V}_b$ can in general be written as the divergent result for the real contribution in $4-2\epsilon$ dimensions with opposite signs in front of the ϵ-poles. We defer the inclusion of massive QCD partons to appendix D.1 and write $\mathcal{V}_b$ in the massless case for a Born flavor structure f_B as [114–116]

$$\begin{aligned}\mathcal{V}_b(\mathbf{\Phi}_n, f_B) &= \mathcal{D}(\epsilon)\frac{\alpha_S}{2\pi}\left[-\sum_{l=n_F}^{n_L^{(B)}+2}\left(\frac{1}{\epsilon^2}C(\bar{\mathcal{I}}_l)+\frac{1}{\epsilon}\gamma(\bar{\mathcal{I}}_l)\right)\mathcal{B}+\frac{1}{2\epsilon}\sum_{\substack{k,l=n_F\\k\neq l}}^{n_L^{(B)}+2}\log\frac{2\bar{k}_k\cdot\bar{k}_l}{Q^2}\mathcal{B}_{kl}+\mathcal{V}_{\text{fin}}\right]_{f_B}^{\mathbf{\Phi}_n}\\ &= \mathcal{D}(\epsilon)\frac{\alpha_S}{2\pi}\left[\sum_{\substack{k,l=n_F\\k\neq l}}^{n_L^{(B)}+2}\left(-\frac{1}{\epsilon^2}+\frac{1}{2\epsilon}\log\frac{2\bar{k}_k\cdot\bar{k}_l}{Q^2}\right)\mathcal{B}_{kl}-\sum_{l=n_F}^{n_L^{(B)}+2}\frac{1}{\epsilon}\gamma(\bar{\mathcal{I}}_l)\mathcal{B}+\mathcal{V}_{\text{fin}}\right]_{f_B}^{\mathbf{\Phi}_n},\end{aligned} \tag{2.3.1}$$

with n_F and $n_L^{(B)}$ defined as in eq. (2.1.10). In the second line we used eq. (2.2.52) to make the ϵ^2-pole proportional to the color-correlated squared Born amplitude $\mathcal{B}_{kl}$. By the super- and subscripts of the square brackets we denote the fact that all quantities inside these brackets are written in terms of the n-body phase space $\mathbf{\Phi}_n$ for a Born flavor structure f_B. The term $\mathcal{V}_{\text{fin}}$ stands for the finite one-loop $\mathcal{O}(\epsilon^0)$-contributions, obtained by the evaluation of loop integrals. In the context of MC event generators it is usually computed by dedicated external programs and the only relevant connection to the FKS subtraction scheme is via the choice of the regularization scheme, where we chose the conventional dimensional regularization (CDR) scheme, and the definition of the normalization factor $\mathcal{D}(\epsilon)$. For the latter we chose the common definition

$$\mathcal{D}(\epsilon) \equiv \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}\left(\frac{\mu^2}{Q^2}\right)^\epsilon, \tag{2.3.2}$$

with the renormalization scale μ_R, where we set $\mu_R = \mu_F = \mu$. We denote by Q the *Ellis-Sexton* scale, which is an arbitrary mass scale introduced in [114] in order to

simplify the result, as Q can be chosen such that logarithms with arguments of Q^2/s or μ^2/Q^2 vanish.

As we have pointed out before, while all ϵ-poles cancel in the sum of all analytically integrated NLO contributions, we need separate finite components for the integration in a MC program. Fortunately, the *virtual-subtracted*[10] squared amplitude for a Born flavor structure f_B, given by [67]

$$\begin{aligned}\mathcal{V}(\mathbf{\Phi}_n, f_B) = \mathcal{V}_b(\mathbf{\Phi}_n, f_B) &+ \left[\sum_{\alpha\in\{\alpha|f_B\}} \Big(I_{s\delta,\alpha} + I_{s+,\alpha} + I_{+\delta,\alpha}\Big)\right]_{f_B}^{\bar{\mathbf{\Phi}}_n^{(\alpha)}=\mathbf{\Phi}_n} \\ &+ \left[\mathcal{G}_{ab,\mathrm{div}}^{(\mathrm{in},\oplus)} + \mathcal{G}_{ab,\mathrm{div}}^{(\mathrm{in},\ominus)}\right]_{f_B}^{\bar{\mathbf{\Phi}}_n^{(\alpha)}=\mathbf{\Phi}_n} , \quad \alpha \in \mathcal{P}_{\mathrm{FKS}} ,\end{aligned} \tag{2.3.3}$$

has no ϵ-poles left and is therefore finite. Here, the integrals I are the ones from eq. (2.2.17), with each of them remaining dependent on the underlying Born kinematics $\bar{\mathbf{\Phi}}_n^{(\alpha)}$ after integrating out the radiation phase space. Obviously, we must only sum over integrals belonging to α-regions (see eq. (2.1.11)) whose underlying Born flavor structure matches f_B. In addition, we keep the underlying Born kinematics $\bar{\mathbf{\Phi}}_n^{(\alpha)}$ fixed and equal to the n-body kinematics of $\mathcal{V}$, denoted by the superscript $\bar{\mathbf{\Phi}}_n^{(\alpha)} = \mathbf{\Phi}_n$. This is also true for the terms $\mathcal{G}_{ab,\mathrm{div}}^{(\mathrm{in},\oplus)}$, which will be derived in section 2.3.3 and the result stated in eq. (2.3.50). They contain the divergent contributions from the initial-state contribution in eq. (2.2.74), which at NLO QCD only contribute for colored initial states. In the following, we will detail the computation of the integrals I.

2.3.1. Integrated Soft Counterterms

We start with the integrated soft counterterms, $I_{s\delta,\alpha}$ and $I_{s+,\alpha}$, and write eq. (2.2.10) as

$$\xi^{-1-2\epsilon} = -\frac{\xi_{\mathrm{c}}^{-2\epsilon}}{2\epsilon}\delta(\xi) + \mathcal{P}_{\mathrm{c}}(\xi) , \quad \mathcal{P}_{\mathrm{c}}(\xi) \equiv \left(\frac{1}{\xi}\right)_{\mathrm{c}} - 2\epsilon\left(\frac{\log\xi}{\xi}\right)_{\mathrm{c}} , \tag{2.3.4}$$

with the same definition for $\mathcal{P}_{\mathrm{c}}$ as in eq. (2.2.70), repeated here for better readability. The soft divergent part of eq. (2.2.17) in $d = 4 - 2\epsilon$ dimensions is then given by

$$I_{s,\alpha} \equiv I_{s\delta,\alpha} + I_{s+,\alpha} = -\frac{s^{1-\epsilon}}{(4\pi)^{d-1}}\frac{\xi_{\mathrm{c}}^{-2\epsilon}}{2\epsilon}\int d\Omega_j^{d-1} d\xi_j \delta(\xi_j)\xi_j^2 \mathcal{R}_\alpha(\xi_j, y_{ij}) , \quad (i,j) \in \mathcal{P}_{\mathrm{FKS}} , \tag{2.3.5}$$

[10] By "subtracted" we refer to the cancellation of the ϵ-poles, not the particular signs of the counterterms.

with the indices i and j of the FKS pair belonging to $\alpha \in \mathcal{P}_{\text{FKS}}$ (see eq. (2.1.11)). We suppress them whenever the context is clear. In eq. (2.3.5), we included the y-dependent part of the phase-space measure in the angular part such that

$$d\Omega^{d-1} = \left(1-y^2\right)^{-\epsilon} dy d\Omega^{d-2} \,. \tag{2.3.6}$$

We take this opportunity to recall that the FKS variables are defined in the CM frame of the incoming partons (see eqs. (2.1.19) and (2.1.21) to (2.1.23)). Consequently, the angular measure is so as well. Due to the delta distribution in ξ_j, we can take the soft limit of $\mathcal{R}_\alpha$, which in $4-2\epsilon$ dimensions is given similarly to eq. (2.2.48) by

$$\mathcal{R}_\alpha(0,y) = \delta_{g\mathcal{I}_j} 4\pi\alpha_S \mu^{2\epsilon} \sum_{k,l=1}^{n+2} \frac{k_k \cdot k_l}{(k_k \cdot k_j)(k_l \cdot k_j)} \mathcal{B}_{kl}(\bar{\Phi}_n^{(\alpha)}) \,, \tag{2.3.7}$$

with the renormalization scale μ and its power 2ϵ stemming from dimensional regularization. By inserting the above equation into eq. (2.3.5) the term $I_{s,\alpha}$ becomes proportional to an ϵ-dependent prefactor times the integral of the eikonal factors such that [41, 42, 59]

$$I_{s,\alpha} = -\frac{\xi_{\text{c}}^{-2\epsilon}}{2\epsilon} \frac{2^{2\epsilon}}{(2\pi)^{1-2\epsilon}} \left(\frac{s}{\mu^2}\right)^{-\epsilon} \frac{\alpha_S}{2\pi} \sum_{k,l=1}^{n+2} \mathcal{B}_{kl}(\bar{\Phi}_n^{(\alpha)}) \int d\Omega_j^{d-1} \left[k_k, k_l\right]_j \,, \tag{2.3.8}$$

where we define

$$\left[k_k, k_l\right]_j \equiv \frac{s\xi_j^2}{4} \frac{k_k \cdot k_l}{(k_k \cdot k_j)(k_l \cdot k_j)} \,, \tag{2.3.9}$$

with the factor ξ_j^2 cancelling the energy dependence of the momenta k_j, making eq. (2.3.8) fully independent of ξ_j. It is common to continue by changing to polar coordinates in order to write the angular phase-space measure as

$$d\Omega^{d-1} = \Omega^{1-2\epsilon} \left(\sin\theta \sin\phi\right)^{-2\epsilon} d\cos\theta d\phi \,, \tag{2.3.10}$$

with the solid angle Ω in $1-2\epsilon$ dimensions given by [101]

$$\Omega^{1-2\epsilon} = 2\frac{(4\pi)^{-\epsilon}\,\Gamma(1-\epsilon)}{\Gamma(1-2\epsilon)} \,. \tag{2.3.11}$$

With the use of the normalization factor $\mathcal{D}(\epsilon)$ given in eq. (2.3.2) and the Taylor expansions

$$\frac{\Gamma(1-\epsilon)\Gamma(1-\epsilon)}{\Gamma(1-2\epsilon)} = 1 - \frac{\pi^2}{6}\epsilon^2 + \mathcal{O}(\epsilon^3) \,, \tag{2.3.12}$$

$$\left(\frac{Q^2}{\xi_c^2 s}\right)^{\epsilon} = 1 + \epsilon \log \frac{Q^2}{\xi_c^2 s} + \frac{\epsilon^2}{2} \log^2 \frac{Q^2}{\xi_c^2 s} + \mathcal{O}(\epsilon^3)\,, \tag{2.3.13}$$

we can rewrite eq. (2.3.8) as

$$\begin{aligned} I_{s,\alpha} = &- \frac{1}{2\epsilon}\mathcal{D}(\epsilon) \left[1 + \epsilon \log \frac{Q^2}{\xi_c^2 s} + \left(\frac{1}{2}\log^2 \frac{Q^2}{\xi_c^2 s} - \frac{\pi^2}{6}\right)\epsilon^2 + \mathcal{O}(\epsilon^3)\right] \\ &\times \frac{\alpha_S}{2\pi} \sum_{k,l=1}^{n+2} \left(\mathcal{B}_{kl}(\bar{\Phi}_n^{(\alpha)}) \int d\cos\theta \frac{d\phi}{\pi} \left(\sin\theta \sin\phi\right)^{-2\epsilon} \left[k_k, k_l\right]_j\right)\,. \end{aligned} \tag{2.3.14}$$

The above integral can then be written in terms proportional to powers of ϵ as [59]

$$\int d\cos\theta \frac{d\phi}{\pi} \left(\sin\theta \sin\phi\right)^{-2\epsilon} \left[k_k, k_l\right]_j = \frac{1}{\epsilon} I_{-1}(k_k, k_l) + I_0(k_k, k_l) + \epsilon I_1(k_k, k_l)\,. \tag{2.3.15}$$

Furthermore, the integrand $\left[k_k, k_l\right]_j$ can be decomposed into a collinear finite and collinear divergent term. The latter term is of the form $1/(1-\cos\theta)$, which gives a factor of $-1/\epsilon$ after integration. The integration of the collinear finite part, however, is rather involved and, besides other terms, gives rise to dilogarithms[11]

$$\mathrm{Li}_2(v) \equiv -\int_0^v \frac{\log(1-u)}{u} du\,, \quad v \in \mathbb{C}\,. \tag{2.3.17}$$

Eventually, the result for the integrated soft counterterm $I_{s,\alpha}$ can be written as [41, 42]

$$I_{s,\alpha} = \frac{\alpha_S}{2\pi} \sum_{k=n_F}^{n_L^{(B)}+2} \sum_{l=k}^{n_L^{(B)}+2} \mathcal{B}_{kl}(\bar{\Phi}_n^{(\alpha)}) \left(\hat{\mathcal{E}}_{kl}^{(m_k,m_l)} + \mathcal{E}_{kl}^{(m_k,m_l)}\right)\,, \tag{2.3.18}$$

where the sum over k and l makes use of the symmetry of $\mathcal{B}_{kl}$. We contained all ϵ-poles due to soft and soft-collinear divergences, as well as all other ϵ dependencies, in the quantity $\hat{\mathcal{E}}_{kl}^{(m_k,m_l)}$. This leaves $\mathcal{E}_{kl}^{(m_k,m_l)}$ finite. Both quantities depend heavily on the masses $m_{k/l}$ of the particles $\bar{\mathcal{I}}_{k/l}$, and have been computed at several occasions with slightly different methods. The massless case can be found in [41], as well as [59]. The latter also deals with the massive-massless and massive cases – as does [100], with the

[11] One very useful identity in this context is given by

$$\mathrm{Li}_2(v) + \mathrm{Li}_2\left(\frac{1}{v}\right) = -\frac{\pi^2}{6} - \frac{1}{2}\log^2(-v)\,. \tag{2.3.16}$$

analytical results published in [42], from where we cite the massless case for $k \neq l$ as

$$\hat{\mathcal{E}}_{kl}^{(0,0)} = \mathcal{D}(\epsilon) \left[\frac{1}{\epsilon^2} - \frac{1}{\epsilon} \log \frac{2\bar{k}_k \cdot \bar{k}_l}{Q^2} + \frac{1}{\epsilon} \log \frac{4\bar{k}_k^0 \bar{k}_l^0}{\xi_c^2 s} \right], \tag{2.3.19}$$

$$\begin{aligned} \mathcal{E}_{kl}^{(0,0)} =& \frac{1}{2} \log^2 \frac{\xi_c^2 s}{Q^2} + \log \frac{\xi_c^2 s}{Q^2} \log \frac{\bar{k}_k \cdot \bar{k}_l}{2\bar{k}_k^0 \bar{k}_l^0} - \mathrm{Li}_2 \left(\frac{\bar{k}_k \cdot \bar{k}_l}{2\bar{k}_k^0 \bar{k}_l^0} \right) \\ &+ \frac{1}{2} \log^2 \frac{\bar{k}_k \cdot \bar{k}_l}{2\bar{k}_k^0 \bar{k}_l^0} - \log \left(1 - \frac{\bar{k}_k \cdot \bar{k}_l}{2\bar{k}_k^0 \bar{k}_l^0} \right) \log \frac{\bar{k}_k \cdot \bar{k}_l}{2\bar{k}_k^0 \bar{k}_l^0} . \end{aligned} \tag{2.3.20}$$

We note that in the soft limit the barred momenta of the set $\bar{\mathbf{\Phi}}_n$ are equal to their counterpart in $\mathbf{\Phi}_{n+1}$. The massless self-eikonals are trivially given by

$$\hat{\mathcal{E}}_{kk}^{(0,0)} = \mathcal{E}_{kk}^{(0,0)} = 0 . \tag{2.3.21}$$

The massive-massless and massive results are given in appendix D.2. It is clear to see that the first two terms of $\hat{\mathcal{E}}_{kl}^{(0,0)}$ cancel the corresponding ϵ-poles in eq. (2.3.1).[12] We will see in section 2.3.3 that part of the last term cancels the corresponding ϵ-poles in the terms $\mathcal{G}_{ab,\mathrm{div}}^{(\mathrm{in},\oplus)}$ of eq. (2.3.3), after using eq. (2.2.52) to rewrite the color-correlated squared Born amplitude in terms of $C(\bar{\mathcal{I}}_l)\mathcal{B}$ and by having $k_\oplus^0 = \sqrt{s}/2$ for ISR. The other part cancels poles occurring in the calculation of the integrated collinear counterterms $I_{+\delta}$, which will be the subject of the following section.

2.3.2. Integrated Final-state Collinear Counterterms

For the calculation of the integrated final-state collinear counterterms, we follow the approach given in [69]. First, we consider the integral[13]

$$\int d\bar{\mathbf{\Phi}}_n^{(\alpha)} I_{+\delta,\alpha} = - \int d\mathbf{\Phi}_{n+1}^{(\alpha)} \frac{\mathcal{P}_c(\xi_j)}{\xi_j^{-1-2\epsilon}} \frac{(2\delta_O)^{-\epsilon}}{\epsilon} \delta(1 - y_{ij})(1 - y_{ij}) \mathcal{R}_\alpha(\xi_j, y_{ij}) , \tag{2.3.22}$$

where we have expanded the right-hand side by $\xi_j^{-1-2\epsilon}$ and in the first step integrate both sides over the n-body phase space only, i.e. we keep the integral over $d\Phi_{\mathrm{rad}}^{(\alpha)}$ at first. The reason for both actions will become clear in the following, as they enable us to rewrite the integrand in favorable ways. For an FKS pair $(i,j) \in \mathcal{P}_{\mathrm{FKS}}$ (see eq. (2.1.11)), we can then factor out the phase space for emitter $\mathcal{I}_i$ and radiated particle

[12] We note that in eq. (2.3.1) we did not make use of the symmetry of $\mathcal{B}_{kl}$, contrary to eq. (2.3.18), thus giving a difference of a factor 2 in corresponding terms.

[13] We note that we defined $I_{+\delta,\alpha}$ without the integration over $\bar{\mathbf{\Phi}}_n$, contrary to [69].

$\mathcal{I}_j$ such that

$$\int d\bar{\mathbf{\Phi}}_n^{(\alpha)} I_{+\delta,\alpha} = -\int dx_\oplus dx_\ominus \, (2\pi)^d \, \delta^d \left(k_\oplus + k_\ominus - \sum_{l=3}^{n+3} k_l \right) \left[\prod_{\substack{l=3 \\ l\neq i,j}}^{n+3} \frac{d^{d-1}k_l}{(2\pi)^{d-1}\, 2k_l^0} \right]$$
$$\times \frac{\left(k_i^0\right)^{1-2\epsilon}}{2\,(2\pi)^{d-1}} dk_i^0 d\Omega_i^{d-1} d\bar{k}_i^0 \delta(\bar{k}_i^0 - k_i^0 - k_j^0) \tag{2.3.23}$$
$$\times \frac{\left(k_j^0\right)^{1-2\epsilon}}{2\,(2\pi)^{d-1}} dk_j^0 dy_{ij} d\Omega_j^{d-2} \frac{\mathcal{P}_{\mathrm{c}}(\xi_j)}{\xi_j^{-1-2\epsilon}} \frac{(2\delta_O)^{-\epsilon}}{\epsilon} \delta\left(1-y_{ij}\right)\left(1-y_{ij}\right) \mathcal{R}_\alpha(\xi_j, y_{ij}) \,,$$

where in the second line we wrote the phase space for the emitter $\mathcal{I}_i$ represented by angular and radial variables, and introduced the energy of the combined momentum $\bar{k}_i = k_i + k_j$. Since the above integral only exists in the collinear limit defined by $\delta(1-y_{ij})$, the momenta have a common direction parametrized by Ω_i. The corresponding energies then are linked by a delta distribution, which we use for the integration over k_i^0. We then obtain the usual replacements (see eq. (2.2.21))

$$k_i^0 = z\bar{k}_i^0 \,, \quad \text{and} \quad k_j^0 = (1-z)\,\bar{k}_i^0 \,. \tag{2.3.24}$$

Together with changing the integration over k_j^0 to an integration over z and using the result for the integration over $d\Omega_j^{d-2}$ with eq. (2.2.3), we obtain

$$\int d\bar{\mathbf{\Phi}}_n^{(\alpha)} I_{+\delta,\alpha} = -\int dx_\oplus dx_\ominus \, (2\pi)^d \, \delta^d \left(k_\oplus + k_\ominus - \sum_{l=3}^{n+3} k_l \right) \left[\prod_{\substack{l=3 \\ l\neq i,j}}^{n+3} \frac{d^{d-1}k_l}{(2\pi)^{d-1}\, 2k_l^0} \right]$$
$$\times \frac{\left(\bar{k}_i^0\right)^{1-2\epsilon}}{2\,(2\pi)^{d-1}} d\bar{k}_i^0 d\Omega_i^{d-1}$$
$$\times \frac{(2\delta_O)^{-\epsilon}}{\epsilon} \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)} \frac{\left(\bar{k}_i^0\right)^{2-2\epsilon}}{8\pi^2} \frac{\mathcal{P}_{\mathrm{c}}(\xi_j)}{\xi_j^{-1-2\epsilon}}$$
$$\times dz z^{1-2\epsilon} \left(1-z\right)^{1-2\epsilon} \lim_{y_{ij}\to 1} \left[\left(1-y_{ij}\right) \mathcal{R}_\alpha(\xi_j, y_{ij}) \right] \,. \tag{2.3.25}$$

The second line then simply is the phase space (in polar coordinates) for the emitter in the underlying Born configuration and, when combined with the first line, it gives the phase space of the underlying Born kinematics $\bar{\mathbf{\Phi}}_n^{(\alpha)}$. We express ξ_j in terms of z such that

$$\xi_j = \xi_{\max}\,(1-z) \,, \quad \text{with} \quad \xi_{\max} = \frac{2\bar{k}_i^0}{\sqrt{s}} \,, \tag{2.3.26}$$

which allows us to write $\mathcal{P}_{\mathrm{c}}$ in terms of z as well, giving us

$$\begin{aligned}\mathcal{P}_{\mathrm{c}}(\xi_j) &= \xi_j^{-1-2\epsilon} + \frac{\xi_{\mathrm{c}}^{-2\epsilon}}{2\epsilon}\delta(\xi_j) \\ &= \xi_{\max}^{-1-2\epsilon}\left[(1-z)^{-1-2\epsilon} + \left(\frac{\xi_{\max}}{\xi_{\mathrm{c}}}\right)^{2\epsilon}\frac{\delta(1-z)}{2\epsilon}\right].\end{aligned} \tag{2.3.27}$$

Furthermore, we already know the collinear limit for the regularized squared real amplitude $(1-z)\,\mathcal{R}_\alpha$ for FSR from eq. (2.2.25), which in $4-2\epsilon$ dimensions is given by

$$\lim_{y_{ij}\to 1}\left[(1-y_{ij})\,\mathcal{R}_\alpha(\xi_j, y_{ij})\right] = \frac{8\pi\alpha_S\mu^{2\epsilon}}{2(\bar{k}_i^0)^2 z\,(1-z)}\langle\hat{P}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}\rangle(z,\epsilon)\mathcal{B}(\bar{\mathbf{\Phi}}_n^{(\alpha)})\,, \tag{2.3.28}$$

where we used eq. (2.2.35) to get rid of the factor $(1-y_{ij})$. As before, the renormalization scale μ and its power of 2ϵ come frome dimensional regularization. Because of the angular integration, which requires the azimuthal average over the $d-2$ transverse components, we can apply the spin-averaged Altarelli-Parisi splitting functions $\langle\hat{P}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}\rangle$ given in eqs. (C.2.2) to (C.2.5). Without the integration over the underlying Born phase space, we have

$$\begin{aligned}I_{+\delta,\alpha} =& -\frac{(2\delta_O)^{-\epsilon}}{\epsilon}\frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}\left(\frac{\mu}{\bar{k}_i^0}\right)^{2\epsilon}\frac{\alpha_S}{2\pi}\int_0^1 dz z^{-2\epsilon}(1-z) \\ &\times\left[(1-z)^{-1-2\epsilon} + \left(\frac{\xi_{\max}}{\xi_{\mathrm{c}}}\right)^{2\epsilon}\frac{\delta(1-z)}{2\epsilon}\right]\langle\hat{P}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}\rangle(z,\epsilon)\mathcal{B}(\bar{\mathbf{\Phi}}_n^{(\alpha)})\end{aligned} \tag{2.3.29}$$

$$\begin{aligned}=& -\frac{1}{\epsilon}\mathcal{D}(\epsilon)\left(\frac{2Q^2}{s\delta_O}\right)^\epsilon\xi_{\max}^{-2\epsilon}\frac{\alpha_S}{2\pi}\int_0^1 dz z^{-2\epsilon}(1-z) \\ &\times\left[(1-z)^{-1-2\epsilon} + \left(\frac{\xi_{\max}}{\xi_{\mathrm{c}}}\right)^{2\epsilon}\frac{\delta(1-z)}{2\epsilon}\right]\langle\hat{P}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}\rangle(z,\epsilon)\mathcal{B}(\bar{\mathbf{\Phi}}_n^{(\alpha)})\,.\end{aligned} \tag{2.3.30}$$

We collected the normalization factor $\mathcal{D}(\epsilon)$ defined in eq. (2.3.2) in the second equality and rewrote

$$\left(\frac{Q^2}{2\delta_O(\bar{k}_i^0)^2}\right)^\epsilon = \left(\frac{2Q^2}{s\delta_O}\right)^\epsilon\xi_{\max}^{-2\epsilon}\,, \tag{2.3.31}$$

using the definition of $\xi_{\max}$ in eq. (2.3.26). The two remaining integrals in eq. (2.3.30) are of the form

$$I^{(0)} \equiv \int_0^1 dz z^{-2\epsilon}(1-z)^{-2\epsilon}\langle\hat{P}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}\rangle(z,\epsilon)\,, \tag{2.3.32}$$

$$I^{(-1)} \equiv \int_0^1 dz z^{-2\epsilon}\left(\frac{\xi_{\max}}{\xi_{\mathrm{c}}}\right)^{2\epsilon}\frac{\delta(1-z)}{2\epsilon}(1-z)\langle\hat{P}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}\rangle(z,\epsilon)\,. \tag{2.3.33}$$

The process of evaluating the integrals $I^{(0)}$ and $I^{(-1)}$ for each kind of splitting function is rather straightforward and has been done e.g. in [69, eqs. (3.48) to (3.50)]. When calculating the integrals, one useful trick is to use the symmetry $z \to 1 - z$ of the integral in the case of a $g \to gg$-splitting. Thus, for this case in particular, the result of eq. (2.3.33) is

$$I_{gg}^{(-1)} = \frac{C_A}{\epsilon} \left(\frac{\xi_{\max}}{\xi_c} \right)^{2\epsilon} . \tag{2.3.34}$$

Eventually, with all integrals calculated, we can write the sum over all α-regions with an associated underlying Born flavor structure f_B for eq. (2.3.30) as

$$\begin{aligned} \sum_{\alpha \in \{\alpha | f_B\}} I_{+\delta,\alpha}(\bar{\mathbf{\Phi}}_n^{(\alpha)}, f_B) = & \sum_{i=n_I+1}^{n_I+n_L^{(B)}} \mathcal{D}(\epsilon) \left(\frac{2Q^2}{s\delta_O} \right)^{\epsilon} \xi_{\max}^{-2\epsilon} \frac{\alpha_S}{2\pi} \mathcal{B}(\bar{\mathbf{\Phi}}_n^{(\alpha)}, f_B) \\ & \times \left[\frac{1}{\epsilon^2} \left(1 - \left(\frac{\xi_{\max}}{\xi_c} \right)^{2\epsilon} \right) C(\bar{\mathcal{I}}_i) + \frac{\gamma(\bar{\mathcal{I}}_i)}{\epsilon} + \gamma'(\bar{\mathcal{I}}_i) \right] \\ = & \, \mathcal{D}(\epsilon) \frac{\alpha_S}{2\pi} \mathcal{B}(\bar{\mathbf{\Phi}}_n^{(\alpha)}, f_B) \sum_{i=n_I+1}^{n_I+n_L^{(B)}} \left(1 - \epsilon \log \left(\frac{s\delta_O}{2Q^2} \right) \right) \\ & \times \left[\frac{\xi_{\max}^{-2\epsilon} - \xi_c^{-2\epsilon}}{\epsilon^2} C(\bar{\mathcal{I}}_i) + \frac{\gamma(\bar{\mathcal{I}}_i)}{\epsilon} \xi_{\max}^{-2\epsilon} + \gamma'(\bar{\mathcal{I}}_i) \xi_{\max}^{-2\epsilon} \right] , \end{aligned} \tag{2.3.35}$$

with n_I and $n_L^{(B)}$ again defined as in eq. (2.1.10). Here, we expanded $(2Q^2/s\delta_O)^{\epsilon}$ in ϵ, and have denoted the Casimir constants $C(\bar{\mathcal{I}}_i)$ as usual. The coefficients of the ϵ^{-1}- and ϵ^0-poles are part of the result of the integration of eq. (2.3.32) and, depending on whether particle $\bar{\mathcal{I}}_i$ is a gluon or (anti-)quark, are given as

$$\begin{aligned} \gamma(g) &= \frac{11}{6} C_A - \frac{2}{3} T_F n_f , & \gamma'(g) &= \left(\frac{67}{9} - \frac{2\pi^2}{3} \right) C_A - \frac{23}{9} T_F n_f , \\ \gamma(q) &= \frac{3}{2} C_F , & \gamma'(q) &= \left(\frac{13}{2} - \frac{2\pi^2}{3} \right) C_F , \end{aligned} \tag{2.3.36}$$

with n_f defined as the number of quark flavors the gluon can split into. Expanding both $\xi_{\max}^{-2\epsilon}$ and $\xi_c^{-2\epsilon}$ in eq. (2.3.35), and collecting the remaining ϵ-poles, gives us the

result

$$\begin{aligned}\sum_{\alpha\in\{\alpha|f_B\}} I_{+\delta,\alpha}(\bar{\mathbf{\Phi}}_n^{(\alpha)}, f_B) =&\mathcal{D}(\epsilon)\frac{\alpha_S}{2\pi}\sum_{i=n_I+1}^{n_I+n_L^{(B)}}\frac{1}{\epsilon}\left[\gamma(\bar{\mathcal{I}}_i) - 2C(\bar{\mathcal{I}}_i)\log\left(\frac{\xi_{\max}}{\xi_c}\right)\right]\mathcal{B}(\bar{\mathbf{\Phi}}_n^{(\alpha)}, f_B)\\ &+\frac{\alpha_S}{2\pi}\sum_{i=n_I+1}^{n_I+n_L^{(B)}}\mathcal{Q}_{+\delta,i}\mathcal{B}(\bar{\mathbf{\Phi}}_n^{(\alpha)}, f_B)\,,\end{aligned} \tag{2.3.37}$$

where we defined

$$\begin{aligned}\mathcal{Q}_{+\delta,i} \equiv&\gamma'(\bar{\mathcal{I}}_i) - \log\left(\frac{s\delta_O}{2Q^2}\right)\left[\gamma(\bar{\mathcal{I}}_i) - 2C(\bar{\mathcal{I}}_i)\log\left(\frac{\xi_{\max,i}}{\xi_c}\right)\right]\\ &+2C(\bar{\mathcal{I}}_i)\left(\log^2\xi_{\max,i} - \log^2\xi_c\right) - 2\gamma(\bar{\mathcal{I}}_i)\log\xi_{\max,i}\,,\end{aligned} \tag{2.3.38}$$

and made the dependence of $\xi_{\max}$ on the momentum of parton $\bar{\mathcal{I}}_i$ explicit by writing the additional index i. The first term in eq. (2.3.37) containing an ϵ-pole cancels the corresponding final-state terms in eq. (2.3.1) that are not canceled by the initial-state terms $\mathcal{G}_{ab,\mathrm{div}}^{(\mathrm{in},\oplus)}$ in eq. (2.3.3) of the following section 2.3.3. The second term containing an ϵ-pole cancels part of the last term in eq. (2.3.19), with the terms $\mathcal{G}_{ab,\mathrm{div}}^{(\mathrm{in},\oplus)}$ again cancelling the remaining part.

2.3.3. Integrated Initial-state Collinear Counterterms and the DGLAP Remnant

We return to the task of integrating the remaining divergent contributions stemming from collinear ISR discussed in section 2.2.2. In this section, we follow the approach given in [41]. The starting point is the integration of the quantity $\mathcal{R}_\alpha^{(\mathrm{in},\oplus)}$, given by eq. (2.2.74), over the radiation phase space in $d = 4 - 2\epsilon$ dimensions such that

$$\int d\Phi_{\mathrm{rad}}^{(\alpha)}\mathcal{R}_\alpha^{(\mathrm{in},\oplus)} = -\frac{s^{1-\epsilon}}{(4\pi)^{d-1}}\frac{(2\delta_I)^{-\epsilon}}{2\epsilon}\int d\Omega^{d-2}d\xi_j dy_j\delta(1\pm y_j)\mathcal{P}_c(\xi_j)\left(1-y_j^2\right)\xi_j^2\mathcal{R}_\alpha(\xi_j, y_j)\,, \tag{2.3.39}$$

with the definition of $\mathcal{P}_c$ given in eq. (2.2.70). The delta distribution in y_j once again allows us to take the collinear limit of the regularized squared real amplitude, given in

the case of ISR as

$$\begin{aligned}\lim_{y_j\to\pm1}\left[\xi_j^2\left(1-y_j^2\right)\mathcal{R}_\alpha(\xi_j,y_j)\right] &= \lim_{y_j\to\pm1}\left[\xi_j^2\left(1-y_j^2\right)\frac{8\pi\alpha_S\mu^{2\epsilon}}{-\bar{k}_i^2}\langle\hat{P}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}\rangle(1-\xi_j,\epsilon)\mathcal{B}(\bar{\mathbf{\Phi}}_n^{(\alpha)})\right]\\ &= \xi_j\frac{8\pi\alpha_S\mu^{2\epsilon}}{(k_i^0)^2}\langle\hat{P}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}\rangle(1-\xi_j,\epsilon)\mathcal{B}(\bar{\mathbf{\Phi}}_n^{(\alpha)})\,,\end{aligned} \tag{2.3.40}$$

where we use the spin-averaged Altarelli-Parisi splitting functions $\langle\hat{P}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}\rangle$, given in eqs. (C.2.2) to (C.2.5), because of the angular integration as explained in the context of eq. (2.3.28). Furthermore, we used eq. (2.2.91) to cancel a factor $(1\mp y_j)$ of $\left(1-y_j^2\right)$, with the remaining one becoming a factor of 2, as well as one order of ξ_j, since we have $\lim_{y\to\pm1} z = 1-\xi$ for ISR (see section 2.2.2). Integrating over the angular measure $d\Omega^{d-2}$, with the result given in eq. (2.2.3), and using the fact that for incoming initial-state momenta we have $k_i^0 = \sqrt{s}/2$, we can insert eq. (2.3.40) into eq. (2.3.39) and get

$$\int d\Phi_{\text{rad}}^{(\alpha)}\mathcal{R}_\alpha^{(\text{in},\oplus)} = -\frac{\alpha_S}{2\pi}\left(\frac{1}{\bar{\epsilon}}-\log\left(\frac{s\delta_I}{2\mu^2}\right)\right)\int d\xi_j\xi_j\mathcal{P}_{\text{c}}(\xi_j)\langle\hat{P}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}\rangle(1-\xi_j,\epsilon)\mathcal{B}(\bar{\mathbf{\Phi}}_n^{(\alpha)})\,, \tag{2.3.41}$$

with $\bar{\epsilon}$ defined in eq. (2.2.63). Here, we once again expanded the prefactor $(2\mu^2/(s\delta_i))^\epsilon$, leaving the term proportional to the resulting logarithm without an ϵ-pole, in which case we can set $\epsilon = 0$. Switching to the shorthand notation for initial-state partons from eq. (2.2.56), we state that in [41] it has been shown that for a *fixed* real flavor structure f_r with an initial-state ab the expression[14]

$$d\hat{\sigma}_{ab}^{(\text{in},\oplus)}(f_r) = d\sigma_{ab}^{(\text{in},\oplus)}(f_r) + d\sigma_{ab}^{(\text{cnt},\oplus)}(f_r) \tag{2.3.42}$$

is finite for purely collinear (i.e. $z<1$) limits. Limiting the PDF counterterms $d\sigma_{ab}^{(\text{cnt},\oplus)}$ to a single fixed real flavor structure eliminates the sum in their definition given in eq. (2.2.67). According to [41], we can write the Altarelli-Parisi kernel $P_{\bar{a}\to a\mathcal{I}_j}(z,0)$ in terms of the spin-averaged splitting functions $\langle\hat{P}_{\bar{a}\to a\mathcal{I}_j}\rangle$. For the $\oplus$-direction it then is

$$P_{\bar{a}\to a\mathcal{I}_j}(z,0) = \frac{(1-z)\,\langle\hat{P}_{\bar{a}\to a\mathcal{I}_j}\rangle(z,0)}{(1-z)_+} + \gamma(\bar{a})\delta_{a\bar{a}}\delta(1-z)\,, \tag{2.3.43}$$

with $\gamma(\bar{a})$ given in eq. (2.3.36). In eq. (2.3.43), we also used the identity

[14] In [41], the second summand has a factor 1/4 which avoids double counting due to the sum over all possible final-state flavors. We do not need this factor here, as we do not perform the flavor sum.

$$\delta(1 \pm y_j) \left(\frac{1}{1-z}\right)_+ = \delta(1 \pm y_j) \left[\left(\frac{1}{\xi_j}\right)_{\mathrm{c}} + \delta(\xi_j) \log \xi_{\mathrm{c}}\right], \tag{2.3.44}$$

where the delta distribution in y_j emphasizes that it is only valid for the collinear limit, as well as the relation

$$2C(\bar{a})\delta_{a\bar{a}}\delta(1-z) = \delta(1-z)\,(1-z)\,\langle \hat{P}_{\bar{a}\to a\mathcal{I}_j}\rangle(z,0)\,, \tag{2.3.45}$$

with the Casimir constants $C(\bar{a})$. For a specific α-region, we can then write the integration over the radiation phase space in $d\sigma_{ab}^{(\mathrm{cnt},\oplus)}(f_r)$ (without the n-body integration measure) with a change of $z \to 1-\xi_j$ as

$$\begin{aligned}\int d\xi_j \mathcal{G}^{\mathrm{cnt},\oplus}_{ab,\mathcal{I}_j}(\mathbf{\Phi}^{(\alpha)}_{n,\oplus}) = \frac{\alpha_S}{2\pi} \int d\xi_j \frac{1}{\bar{\epsilon}} \Bigg[&\left(\frac{1}{\xi_j}\right)_{\mathrm{c}} \xi_j \langle \hat{P}_{\bar{a}\to a\mathcal{I}_j}\rangle(1-\xi_j,0) \\ &+ \delta_{a\bar{a}}\delta(\xi_j)\,(\gamma(\bar{a}) + 2C(\bar{a})\log\xi_{\mathrm{c}})\Bigg] \mathcal{B}_{\bar{a}b}(\bar{\mathbf{\Phi}}^{(\alpha)}_n)\,.\end{aligned} \tag{2.3.46}$$

Here, we explicitly wrote the initial-state flavors as subscripts for the squared Born amplitude for clarification. The $\ominus$-direction is of course completely analogous for the splitting $\bar{b} \to b\mathcal{I}_j$. By expanding

$$\langle \hat{P}_{\bar{a}\to a\mathcal{I}_j}\rangle(1-\xi_j,\epsilon) = \langle \hat{P}_{\bar{a}\to a\mathcal{I}_j}\rangle(1-\xi_j,0) + \epsilon \left.\frac{\partial \langle \hat{P}_{\bar{a}\to a\mathcal{I}_j}\rangle(1-\xi_j,\epsilon)}{\partial\epsilon}\right|_{\epsilon=0} + \mathcal{O}(\alpha_S^2)\,, \tag{2.3.47}$$

and adding eqs. (2.3.41) and (2.3.46), as suggested by eq. (2.3.42), we get the expressions

$$\begin{aligned}\mathcal{G}^{(\mathrm{in},\oplus)}_{ab,\mathcal{I}_j}(\mathbf{\Phi}^{(\alpha)}_{n,\oplus}) = \frac{\alpha_S}{2\pi} \Bigg\{ &(1-z)\,\langle \hat{P}_{\bar{a}\to a\mathcal{I}_j}\rangle(z,0) \left[\left(\frac{1}{1-z}\right)_{\mathrm{c}} \log\frac{s\delta_I}{2\mu^2} + 2\left(\frac{\log(1-z)}{1-z}\right)_{\mathrm{c}}\right] \\ &- (1-z) \left.\frac{\partial \langle \hat{P}_{\bar{a}\to a\mathcal{I}_j}\rangle(z,\epsilon)}{\partial\epsilon}\right|_{\epsilon=0} \left(\frac{1}{1-z}\right)_{\mathrm{c}} \Bigg\} \mathcal{B}_{\bar{a}b}(\bar{\mathbf{\Phi}}^{(\alpha)}_n) + \mathcal{G}^{(\mathrm{in},\oplus)}_{ab,\mathrm{div}}(\bar{\mathbf{\Phi}}^{(\alpha)}_n)\,,\end{aligned} \tag{2.3.48}$$

$$\begin{aligned}\mathcal{G}^{(\mathrm{in},\ominus)}_{ab,\mathcal{I}_j}(\mathbf{\Phi}^{(\alpha)}_{n,\ominus}) = \frac{\alpha_S}{2\pi} \Bigg\{ &(1-z)\,\langle \hat{P}_{\bar{b}\to b\mathcal{I}_j}\rangle(z,0) \left[\left(\frac{1}{1-z}\right)_{\mathrm{c}} \log\frac{s\delta_I}{2\mu^2} + 2\left(\frac{\log(1-z)}{1-z}\right)_{\mathrm{c}}\right] \\ &- (1-z) \left.\frac{\partial \langle \hat{P}_{\bar{b}\to b\mathcal{I}_j}\rangle(z,\epsilon)}{\partial\epsilon}\right|_{\epsilon=0} \left(\frac{1}{1-z}\right)_{\mathrm{c}} \Bigg\} \mathcal{B}_{a\bar{b}}(\bar{\mathbf{\Phi}}^{(\alpha)}_n) + \mathcal{G}^{(\mathrm{in},\ominus)}_{ab,\mathrm{div}}(\bar{\mathbf{\Phi}}^{(\alpha)}_n)\,.\end{aligned} \tag{2.3.49}$$

where we defined

$$\mathcal{G}^{(\text{in},\oplus)}_{ab,\text{div}}(\bar{\mathbf{\Phi}}^{(\alpha)}_n) \equiv \frac{\alpha_S}{2\pi}\left[\mathcal{D}(\epsilon)\frac{1}{\epsilon} - \log\frac{\mu^2}{Q^2}\right]\left(\gamma(\bar{\mathcal{I}}_\oplus) + 2C(\bar{\mathcal{I}}_\oplus)\log\xi_\text{c}\right)\mathcal{B}_{ab}(\bar{\mathbf{\Phi}}^{(\alpha)}_n)\,. \quad (2.3.50)$$

Here, we simply added and subtracted a term proportional to $(\mu^2/Q^2)^\epsilon$ in order to write the normalization factor $\mathcal{D}(\epsilon)$ as defined in eq. (2.3.2). The term $\mathcal{G}^{(\text{in},\oplus)}_{ab,\text{div}}$ still has an explicit ϵ-pole, which can be associated with a soft divergence. As stated at the end of sections 2.3.1 and 2.3.2, however, it is canceled by corresponding terms in the virtual-subtracted squared amplitude given in eq. (2.3.3). The rest of $\mathcal{G}^{(\text{in},\oplus)}_{ab,\mathcal{I}_j}$ is finite and can thus be integrated as an additional NLO contribution. Therefore, it is sometimes referred to as the *DGLAP remnant*. Another way to refer to the remnant, present in the literature [42], is to call it the *degenerate* $(n+1)$*-body kinematics*. Why this is a fitting description can be understood if we look at the kinematics and the resulting phase space for this contribution, which is already discussed in section 2.2.2.

2.3.4. The Complete Virtual-subtracted Squared Amplitude

We have shown that all divergences made explicit by ϵ-poles in $d = 4-2\epsilon$ dimensions in the squared bare one-loop amplitude $\mathcal{V}_b$, given in eq. (2.3.1), cancel against corresponding poles in the integrated counterterms. The virtual-subtracted squared amplitude $\mathcal{V}$, given in eq. (2.3.3), can now be written in a very similar form to $\mathcal{V}_b$ such that

$$\mathcal{V}(\mathbf{\Phi}_n, f_B) = \frac{\alpha_S}{2\pi}\sum_{k=n_F}^{n_L^{(B)}+2}\sum_{l=k}^{n_L^{(B)}+2}\mathcal{E}^{(m_k,m_l)}_{kl}\mathcal{B}_{kl}(\mathbf{\Phi}_n, f_B) + \frac{\alpha_S}{2\pi}\sum_{k=n_I+1}^{n_L^{(B)}+2}\mathcal{Q}_k\mathcal{B}(\mathbf{\Phi}_n, f_B) + \mathcal{V}_\text{fin}(\mathbf{\Phi}_n, f_B)\,, \quad (2.3.51)$$

with the notation of the lower and upper bounds in the sum over the indices k and l of the n-body kinematics $\mathbf{\Phi}_n$ as detailed in eq. (2.1.10). Here, we use $\mathbf{\Phi}_n$ instead of $\bar{\mathbf{\Phi}}_n$ for reasons stated in the context of eq. (2.3.3). The factor $\mathcal{Q}_k$ is defined as

$$\mathcal{Q}_k \equiv \mathcal{Q}_{+\delta,k} + \log\frac{\mu^2}{Q^2}\left[\gamma(\bar{\mathcal{I}}_\oplus) + \gamma(\bar{\mathcal{I}}_\ominus) + 2\big(C(\bar{\mathcal{I}}_\oplus) + C(\bar{\mathcal{I}}_\ominus)\big)\log\xi_\text{c}\right]\,, \quad (2.3.52)$$

with $\mathcal{Q}_{+\delta,k}$ given by eq. (2.3.38). The remaining contribution to $\mathcal{Q}_k$ is due to the finite part of the terms $\mathcal{G}^{(\text{in},\oplus)}_{ab,\text{div}}$, which contain the divergent contributions from the initial-state contribution, given by eq. (2.3.50).

Additionally, in eq. (2.3.51), the finite part $\mathcal{E}^{m_k,m_l}_{kl}$ of the eikonal integrals are given in eq. (2.3.20) for the massless case, with the massless-massive and massive results given in appendix D.2. For the latter cases, we have to adjust the sum to include the number of heavy QCD particles n_H as well. Last, but not least, we have the finite

one-loop $\mathcal{O}(\epsilon^0)$-contributions $\mathcal{V}_{\text{fin}}$ from $\mathcal{V}_b$, which do not have to be subtracted.

The integrated virtual-subtracted cross section for a Born flavor structure f_B can then finally be written as

$$\sigma^{\text{virt}}_{\text{NLO}}(f_B) = \frac{J^{n_L^{(B)}}}{\mathcal{N}(f_B)} \int d\mathbf{\Phi}_n \mathcal{V}(\mathbf{\Phi}_n, f_B)\,, \tag{2.3.53}$$

with the condition $J^{n_L^{(B)}}$ explained around eq. (2.1.7). The symmetry factor $\mathcal{N}(f_B)$ accounts for the presence of identical final-state particles.

3. The Monte Carlo Event Generator WHIZARD at NLO QCD

The multi-purpose MC event generator WHIZARD [35, 98, 117–120] offers a broad range of features desirable for phenomenological HEP studies, such as the computation of cross sections, differential distributions, and fully exclusive event samples of HEP processes. These features are supported for scattering processes at particle colliders with leptonic, hadronic, and mixed leptonic-hadronic initial states. Additionally, particle-decay processes are supported as well.

WHIZARD has roots deep in the studies of lepton colliders of both past and future, as it was originally developed in the context of technical design studies [121, 122] for the planned, but eventually discarded, 800 GeV lepton collider TESLA. There, the need arose to handle more complicated processes such as vector-boson scattering with mixed fermionic and bosonic final states with up to six particles. The intended purpose of WHIZARD was to address this need and the deprecated acronym of the name "*W*, *Hi*ggs, *Z*, *A*nd *R*espective *D*ecays" still reminds of this origin. Since then, WHIZARD has been established as an important tool for a vast amount of lepton-collider studies, as well as design reports [22, 123–125], and has undergone significant expansions in its features.

So far, the program had two major revisions. The first was from WHIZARD 1 to WHIZARD 2. This included an extensive rewriting of the whole program, which led to its current modern, object-orientated, and modular structure, combining several purpose-build packages and interfaces into one coherent and fully automated process. It is written almost entirely in modern `Fortran` – i.e. based on the Fortran 2008 standard – making use of its ubiquitous advantages in numerical computations. The second revision to its current version WHIZARD 3.0.0 has been released just before the completion of this thesis. It marks the first official support of fully automated NLO QCD computations in the WHIZARD series, based upon work associated with this thesis and the prior limited implementation by two former PhD studies [26, 73].

WHIZARD comes with its own matrix-element generator O'Mega [126], which

computes matrix elements via a recursive algorithm using helicity amplitudes and it treats QCD in the color-flow formalism [127]. It is written in `O'Caml` and generates tree-level matrix-element code in `Fortran` at run time, either as compiled libraries or as byte-code instructions interpretable by the O'Mega virtual machine [128], for multi-leg processes in the SM and beyond. This provides a high level of flexibility, since WHIZARD does not need a curated library of precomputed processes as a result.

Given its modularity, WHIZARD supports interfaces to other matrix-element generators as well, both at one-loop and tree-level. This is an important feature especially at NLO, since for the time being O'Mega can neither provide one-loop, nor spin- and color-correlated Born matrix elements, which are needed in the collinear and soft approximations of real matrix elements (see eqs. (2.2.26) and (2.2.51)). Thus, this task is outsourced to third-party OLPs, namely OpenLoops 2 [52, 53] (and its third party tools [44–47]), GoSam [48], and RECOLA [54–57]. The first two are interfaced via the Binoth Les Houches Accord (BLHA) interface [129, 130], whereas RECOLA provides its own dedicated interface and respective implementation.

However, as of WHIZARD 3.0.0, only the interface to the matrix-element generator and OLP OpenLoops is fully complete for NLO in the context of FKS. This will change in the near future with the planned capability of O'Mega to provide spin- and color-correlated Born matrix elements. Consequently, O'Mega will then be able to provide all parts needed for the NLO computation, except for the virtual one-loop amplitude, which can already be obtained via the RECOLA and GoSam interfaces.

Phase-space integration in WHIZARD is done in a multi-channel twofold-adaptive[1] approach via an amplified implementation of the Vegas [131, 132] algorithm, provided by the Vamp [133] module. A modernized and improved version is given with Vamp2 [134, 135], which provides a modern message passing interface (MPI) based parallelization, where a thread-safe random number generator (RNG) is employed.

Another very distinctive feature of WHIZARD is its ability to describe and provide various elements of lepton-collider physics. It is able to properly simulate polarized beams, crossing angles, and photon-induced background processes. Electron PDFs are supported, with soft photons resummed to all orders of the QED coupling α, while hard-collinear photon resummation is supported up to third order. ISR photon-p_T spectra generation and lepton-collider beam-spectra sampling are provided within the WHIZARD framework by the lepton-collider beam-spectrum generator Circe1 [136] and its modernized and improved version Circe2.

While the feature suite of WHIZARD is optimized primarily for lepton-collider physics, it supports hadronic beams as well. For this purpose, a small set of built-in PDF sets are provided, with the option to interface LHAPDF 6 [137] to access a

[1] Meaning adaptive both in the grid and the event weights. See sections 3.1.2 and 3.1.5 for more details.

broad variety of additional PDF sets.

Effects of FSR, like parton showers or hadronization, can be handled in various ways. WHIZARD comes with its own analytic shower [138], as well as with the PYTHIA6 [36] library included in the WHIZARD package for showering and hadronization. Furthermore, PYTHIA8 [37] is directly supported via an interface [139].

In addition, any external simulation or analysis tool can be interfaced, provided it supports any of the numerous event formats that WHIZARD can output to. These are LCIO [140], HEPMC2&3 [141, 142], LHEF [143–145], and several others, including different ASCII file formats. Furthermore, WHIZARD itself can be used as a library via the recently implemented Application Programming Interface (API), which – beside others – includes a Python interface.

WHIZARD has always had a strong focus on BSM physics and has been used for many studies in this field [146–157]. Internally, each process is linked to a certain physics model, either provided by the multitude of predefined models shipped directly with WHIZARD, or provided via automated interfaces to the external tools FEYNRULES [158, 159] or SARAH [160]. Recently, these interfaces have been superseded with the support of the universal FEYNRULES output (UFO) format [161], enabling the user to compute processes in almost any arbitrary, Lagrangian-based BSM theory at LO.[2]

A very powerful tool to steer the whole of WHIZARD is given by its domain-specific language SINDARIN (*S*cripting *in*tegration, *d*ata *a*nalysis, *r*esults display, and *in*terfaces). It supports settings such as process and beam definitions, parameter specifications and model selections. Contrary to most other programs using some form of parameter and/or run cards with static input, SINDARIN enables the user to dynamically set cuts, scales, and scans. This is possible not only for integration and event generation, but also for the integrated analysis suite, where histograms and plots can be defined. In order to provide this flexibility, SINDARIN features common structures of programming languages such as conditional constructs and loops. For jet clustering, the interface to FASTJET [162] is directly steerable from within SINDARIN as well. We provide an example SINDARIN input file in appendix F.1.

In the following, we provide an overview of the aspects of WHIZARD that are relevant in the context of automated NLO computations and highlight the parts that had to be changed or improved in order to support arbitrary processes at fNLO QCD.

[2] BSM models that do not change aspects of QCD could in principle be supported at NLO QCD via the FKS implementation in WHIZARD. However, this remains to be validated.

3.1. WHIZARD in the Context of NLO Calculations

Earlier applications of WHIZARD in the context of NLO calculations have been chargino production at the ILC [163, 164], in which case soft-photon resummation and fixed-order QED and EW corrections have been studied. Furthermore, the process $pp \to b\bar{b}b\bar{b}$ at NLO QCD [165, 166] has been studied with WHIZARD using the CS subtraction scheme [40]. These efforts, however, have been process specific and did not lead to a fully automated implementation. For NLO QCD, a first proof of concept of such an automation in WHIZARD based on FKS subtraction has been given in [97], where its functionality has been shown for off-shell top quark production at a lepton collider. These initial efforts were followed by more thorough studies of top physics at NLO QCD at lepton colliders, both in the continuum [167] and in the threshold region [99].

This initial implementation of the FKS subtraction scheme has paved the way towards fully automated NLO QCD computations for lepton and hadron colliders in WHIZARD. So far, it has been validated for a number of lepton-collider processes limited to FSR of a gluon [26]. More complicated processes with arbitrary external partons were not possible within the early implementation. Furthermore, besides some first ad hoc tests for a Drell-Yan process with a single quark flavor, no hadron-collider processes were supported.

With the work introduced as part of this thesis, WHIZARD is capable of computing completely general external states for lepton- and hadron-collider NLO QCD processes at FO. In order to accomplish this, it was necessary to rework several key aspects of the (NLO) infrastructure of WHIZARD.

The implementation details include notable changes in how WHIZARD keeps track of the variety of different squared amplitudes needed for an NLO calculation and how they are mapped to the correct flavor structures, which is outlined in more detail in section 3.1.1. There, we also give an overview of the general internal structure of WHIZARD and how it organizes and combines the necessary pieces of a computation in the context of NLO.

A brief introduction to the topic of MC phase-space integration is given in section 3.1.2, where we also present the phase-space parametrization realized in WHIZARD in the context of NLO.

For the generalization of NLO QCD computations, necessary changes were introduced in the setup of the singular-region data for the correct application of the $\mathcal{S}$-functions (see section 2.1.2) in order to ensure the correct partitioning of the phase space. A more detailed explanation of how this is realized is given in section 3.1.3.

Furthermore, the correct symmetry factors have to be applied for identical particles in the final state. This is detailed in section 3.1.4. There, we also explain our generalized approach of applying the correct (rescaled) structure functions to the real counterterms and the DGLAP remnant. In order to fully support general NLO QCD hadron-collider processes, the implementation of the DGLAP remnant had to be rewritten almost completely, where we spare the reader the very technical details.

Finally, as a vital step towards parton shower matching, fixed-order differential distributions had to be partly reimplemented and validated for completely general differential distributions in arbitrary processes. In section 3.1.5, we give an overview of event generation at NLO in general, together with the details of the implementation of event generation in WHIZARD at fNLO QCD.

3.1.1. The General Structure of FKS Terms

The internal structure of WHIZARD that organizes and combines the necessary pieces of an NLO computation is rather straightforward and inspired by the terms and squared matrix elements present in the FKS subtraction scheme, discussed in chapter 2. Also, many of the concepts presented in the following are built upon existing LO infrastructure within WHIZARD that is adapted for NLO calculations. For example, at LO WHIZARD can directly compute results for the user-defined sum of different processes, such as $e^+e^- \to ZH$ and $e^+e^- \to ZZ$, which require different phase-space setups internally. Thus, calculations in WHIZARD are structured into different components and terms for the different processes. Further into this section, we explain how this structure is adapted for NLO calculations, where we have different kinematics for the n-body and $(n+1)$-body phase space. First, however, we begin with the organization of matrix elements and subsequently work our way up in the structural hierarchy.

Data Types of State Matrices, Interactions, and Evaluators

While the number of matrix elements for less complex processes – e.g. processes that have a small number of contributing Feynman graphs or low multiplicities – are quite manageable, the number of terms in the amplitude can become quite large for more complex processes. This is especially the case for matrix elements generated by O'Mega, as they are exclusive in the flavors f_i, the color flow indices c_i [127], and the helicities h_i of the process, with i denoting the position of the particle associated with these quantum numbers in a flavor structure (see eq. (2.1.9)). Each set $\{f_i, c_i, h_i\}$

is represented in WHIZARD as a `quantum_numbers_t` type.[3] In total, for matrix elements fully exclusive in all of the mentioned quantum numbers, the dimension of storage required would be $n_f \cdot n_c \cdot n_h$. It becomes obvious that we need a clear, easy to traverse, and dynamically modifiable data structure – ideally with established, fast algorithms for operations such as searches, and insertion or deletion of data.

In WHIZARD, this is realised through the implementation of a so-called *trie*, an ordered tree data structure with nodes, each given a set of quantum numbers $\{f_i, c_i, h_i\}$. Therefore, moving along connected nodes leads to a sequence of quantum-number sets, defined by the position of the last encountered node. Each full sequence of quantum number sets at the end of a node sequence represents a single flavor structure of the associated process and is assigned a matrix element `ME(i)` with integer index `i`. We call the implementation of such a trie in WHIZARD a *state matrix* in direct correspondence to its physical meaning.

In a subtraction scheme like FKS, however, we often require more than one (squared) matrix element per flavor structure. Explicitly, we additionally need

- for the real counterterms
  - 6 entries per emitter $\mathcal{I}_i$ for the independent, non-trivial spin-correlated squared Born matrix elements $\mathcal{B}^{(i)}_{\mu\nu}$ (see eq. (2.2.26)),[4]
  - $n_S\,(n_S-1)\,/2$ entries for the independent color-correlated squared Born matrix elements $\mathcal{B}_{kl}$ (see eq. (2.2.51), with n_S defined in eq. (2.1.10)),
- for the integrated counterterms (see eq. (2.3.51))
  - an entry for the finite one-loop contribution $\mathcal{V}_{\text{fin}}$,
  - $n_S\,(n_S-1)\,/2$ entries for the independent color-correlated squared Born matrix elements $\mathcal{B}_{kl}$,
- and an entry for each incoming particle to account for different squared matrix elements resulting from rescaling its momentum fraction due to collinear ISR, accounted for in the real counterterms and the DGLAP remnant (see sections 2.2.2 and 2.3.3).

In order to account for these entries in the state matrix, we simply introduce an additional pseudo quantum number $\sigma_i \in [0, n_{\text{add}}]$, ranging from 0 (tree-level) to the total number n_{add} of additionally required entries. The original state matrix with its quantum numbers is then continued by n_{add} copies of its structure, each with a

[3] In object-orientated programming, we generally speak of a *class* in order to refer to user-defined data structures with associated methods. In `Fortran` this is called a user-defined derived-type, or *type* for short, whereas a `Fortran` *class* describes a type and any of its possible extensions.

[4] We choose $k_\perp = (0, \mathbf{k}_\perp)$ in eqs. (2.2.23) and (2.2.81), leading to vanishing terms in $\mathcal{B}^{(i)}_{\mu\nu}$.

different σ_i. An example for a state matrix including the discussed attributes so far is shown in fig. 3.1.

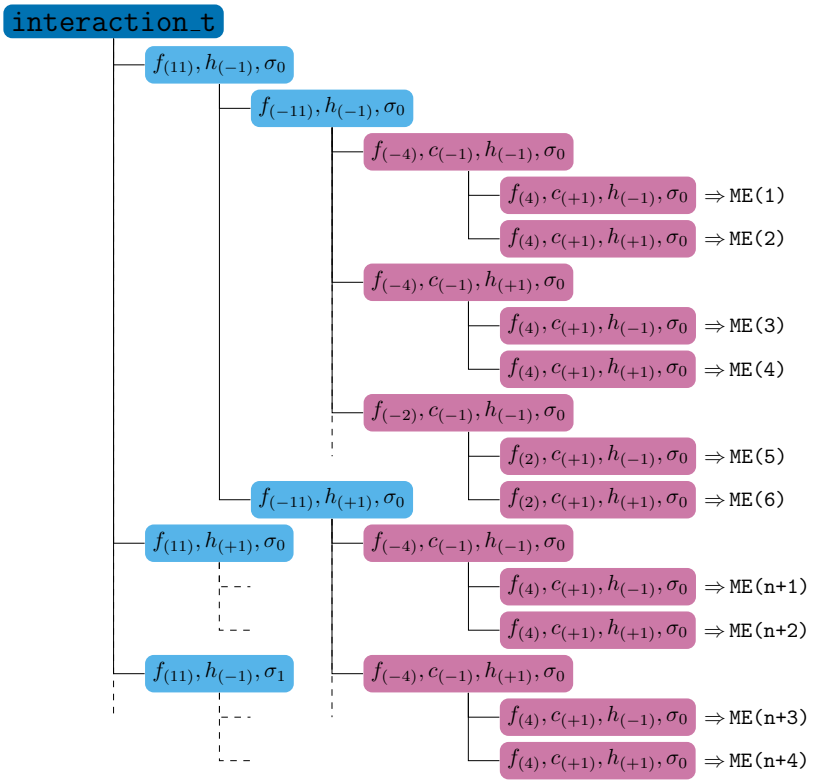


Figure 3.1. Representation of a state matrix trie for the process $e^+e^- \to jj$, for the purpose of demonstration given fully exclusive in the quantum numbers $\{f_i, c_i, h_i\}$ and the pseudo quantum number σ_i, as explained in the text. We only show the first necessary nodes that are required to infer the rest of the state matrix structure, implied by the dashed lines. The number `n` in `ME(i)` simply represents the total numbers of final state configurations for the first initial state configuration at the top. The first node, the root of the trie, is associated with an empty set of quantum numbers. Here, we labelled it to emphasize the relation between the `state_matrix_t` type and the `interaction_t` type it is a member of.

As mentioned before, O'Mega provides matrix elements exclusive in color and helicity. However, at NLO we currently employ external matrix-element generators (e.g. OpenLoops) which mostly provide squared amplitudes inclusive – i.e. summed over – (final state) color and helicity. In order to account for this fact, WHIZARD employs so-called quantum number masks that selectively act on the quantum numbers of the state-matrix nodes to contract the state matrix in the way defined by the mask. For example, if we want to include beam polarization, a mask can be configured to only contract final-state helicities, leaving the initial-state helicity structure explicit.

Depending on the process definition and multiplicity, state matrices can get rather large. However, in the context of the FKS implementation, we rarely make use of the

full quantum-numbers information for each particle in a flavor structure. Throughout the FKS-specific code, we often use simpler arrays of Particle Data Group (PDG) [168] code sequences. These are stored in arrays with an integer index `i_flv` for a certain flavor structure, and, when needed, with an additional integer index `i_hel` to differentiate between equal flavor structures with different helicity configurations.

In order to correctly interconnect the (squared) matrix elements in the state matrix and the simpler flavor array, we needed to implement a mapping between the flavor-structure indices and the corresponding matrix-element integer index `i`, especially since in general the ordering of the flavor array is different than that of its representation in the state matrix. In addition, to get the correct kind of matrix element – i.e. tree-level, spin- or color-correlated – we also have to map the pseudo quantum number σ_i. This is now all handled via the `qn_index_mapping_t` type.

In WHIZARD, both a state matrix and its respective `qn_index_mapping_t` are a member of the `interaction_t` type. The interaction contains additional data, such as phase-space and particle information, and the aforementioned quantum number masks. Two different interactions can be connected in different ways by the `evaluator_t` type, which inherits from `interaction_t` in the sense of object-oriented programming.

How exactly the two interactions are convolved is handled by so-called pairing arrays, which are members of the evaluators. Each pairing array defines the convolution rules between two interactions by acting on the indices of their state matrices. The correct helicity- and color-averaged sum of the resulting amplitudes is handled then as well.

A prime example for the use of an evaluator is in the case of processes that require structure functions, such as PDFs for hadronic beams or QED ISR structure functions for leptonic beams. In WHIZARD, any kind of structure function or beam spectrum is represented as a type of interaction, where the corresponding state matrix only covers the initial state. For example, in the case of leptonic beams the need to concatenate a beamstrahlung spectrum with structure functions exists. They can be conveniently chained together with the evaluator type, which results in an effective structure function interaction that can then be convolved with the hard interaction. The evaluators for this last step are members of the `connected_state_t` type and explicitly are

- the trace evaluator, that sums over all (non-pseudo) quantum numbers,
- the transition-matrix evaluator, that sums over all unobservable quantum numbers,
- and one evaluator for the transition matrix without interferences and exclusive in color flow.

The trace is mostly needed in the computation of integrated cross sections, the transition matrix is relevant for event generation, and the color flows are important in

the subsequent matching to a parton shower.

Term- and Process-Instance Data Types

WHIZARD allows the user to define an LO process as the sum of processes. Internally, this is handled by creating one process that is divided into components in order to handle differences of the processes in the sum – e.g. different phase-space setups due to differing masses of the particles participating in the process, different multiplicities, etc. The underlying structure is adopted for NLO computations where we separate the calculation of the cross section into different components of the FKS subtraction scheme as introduced in chapter 2. In summary, these components are

- the Born component, with n-body kinematics,
- the real component, with $(n+1)$-body kinematics for the non-subtracted contribution and effective n-body kinematics for the subtraction terms (see section 2.2),
- the virtual component, with n-body kinematics, containing the finite one-loop contribution and the integrated counterterms (c.f section 2.3),
- the DGLAP remnant, with degenerate $(n+1)$-body kinematics (see section 2.3.3),
- specialized optional NLO components, e.g. the real-finite or soft mismatch (see text below).

Additional components, together with their own kinematics, are supported by WHIZARD as well. For example, these can be the component for the computation of the finite tree-level real amplitude used for `POWHEG` damping [59] or the soft mismatch, the latter of which being an additional component in the resonance-aware expansion of the FKS subtraction scheme [69]. The implementations for both components in WHIZARD are documented in [26, 73].

As stated in the list above, different NLO components require different kinematics, process and particle information, evaluators, etc. We can even have different sets of kinematics inside a single component as is the case for the real component.[5] Also, each component requires its own provision and evaluation of (subtracted) amplitudes, as described in chapter 2.

In order to store this information and to give each component the corresponding content, as well as to calculate the correct amplitude, WHIZARD initializes so-called

[5] To be technically correct, this is handled by an additional "pseudo-component" for the subtraction terms that is needed for the internal WHIZARD infrastructure, but is marked as inactive as it does not carry out its own integration. This is because we have to integrate the sum of the terms in eq. (2.2.20) over the full real phase space.

`term_instance_t` types, one for each independent kinematic configuration (as classified according to FKS). This means that we have one term instance for each component in the list above, except for the real component. It has $(n_{\text{em}} + 1)$ instances, since we have as many independent real-emission kinematic configurations as the number n_{em} of unique emitters (see section 3.1.3) plus one Born configuration for the full subtraction term.

Term instances can be set active or inactive and are consequently included or excluded in the program evaluation, respectively, allowing for more flexibility in the computation as we will discuss at the end of section 3.1.2. Each term instance holds the interaction type for the hard process, as well as the necessary evaluators for the evaluation of the convolution of structure functions with the hard interaction.

Components and their respective term instances are part of the `process_instance_t` type. As the name suggests, this type represents the full process and contains all necessary top-level information to compute the full amplitude, and to then pass it on to the corresponding multi-channel integrator (MCI) (e.g. VAMP`1/2`).

The process-component manager, of type `pcm_t`, handles all the different components inside the process instance, its MCIs, etc. Its NLO extension includes further data needed for the FKS implementation such as the singular-region data, discussed in more detail in section 3.1.3. The `pcm_t` type also handles the assignment of the different matrix-element generators using a dedicated structure which we refer to as the matrix-element *core*. In principle, the matrix-element generators might differ between the different term instances. For example, we could obtain the spin- and color-correlated squared amplitudes, as well as the finite one-loop contributions, from OpenLoops, while employing O'Mega for the remaining amplitudes. Thus, every term instance has a pointer to a certain core it might share with another term instance.

The evaluation of (squared) matrix elements by a matrix-element generator can be very time-consuming, especially for one-loop contributions. For those matrix-element generators that support the BLHA interface such as OpenLoops, we optimized calls from the term instance to the core in such a way that we only request squared matrix elements for non-equivalent subprocesses. For example, in the case of the process $e^+e^- \to q\bar{q}$ with massless quarks, the subprocesses $e^+e^- \to u\bar{u}$ and $e^+e^- \to c\bar{c}$ are equal with respect to their squared amplitude, while $e^+e^- \to d\bar{d}$ gives a different result. The selection which subprocesses are equivalent with respect to their squared amplitude is possible since the information is provided by the matrix-element generator via the BLHA interface. For processes that have a large number of equivalent subprocesses (such as $e^+e^- \to jjjjjj$, see chapter 5) we can more than halve the computation time.

After evaluating each active term instance, its matrix and color-flows evaluators are filled with the corresponding computed exclusive amplitudes. The inclusive sum, however, is handed over by the process instance to the MCI.

3.1.2. Phase-space Parametrization and Integration

In particle physics, we are typically faced with multidimensional integrals that have not only complicated integrands in the form of squared amplitudes, but also exhibit regions of phase space governed by kinematic effects such as resonances, or radiation peaks, leading to a complex structure within the phase space.

In general, following [134], we can write the squared amplitude as a function $f : \Pi \to \mathbb{R}$, mapping the compact phase-space manifold Π of d dimensions to a real-valued scalar. The phase-space integral then is defined as

$$I_\Pi[f] \equiv \int_\Pi d\mu(p) f(p)\,, \tag{3.1.1}$$

where we denote by p not only four-momenta, but also every other possible integration variables that might occur, e.g. due to structure functions such as PDFs. Furthermore, we assume Π and the integration measure $d\mu$ to account for momentum conservation and on-shell conditions, as well as optional constraints like user-defined cuts. In an MC computation, for the generation of p we use a set x of random numbers $x_i \in U$ represented by the subset $U \subset (0,1)^d$ of the unit hypercube in d dimensions. The phase space is then parametrized by a bijective mapping $\varphi : U \to \Pi$ such that

$$p = \varphi(x)\,, \quad d\mu(p) = \varphi'(x) d\mu(x) = \varphi'(x)\rho_\varphi(x) d^d x\,, \tag{3.1.2}$$

with the Jacobian $\varphi' = \det(d\varphi/dx)$. The density $\varphi'(x)\rho_\varphi(x)$ thus relates the measure $d^d x$ on $\mathbb{R}^d$, expressed in its canonical basis, to the phase-space measure $d\mu$ [134].

Since the Feynman diagrams of a process can be directly related to the underlying kinematic structure, for each scattering (sub)process WHIZARD defines the phase-space parametrizations recursively in a tree structure, similar to Feynman graphs. Here, however, we treat the process as if there was an initial particle that subsequently decays via nodes that can either split into two daughter particles or do not have any further branches, thus representing possible flows of momenta in a process. Resonances and other kinematically challenging regions are being kept track of within these trees.

Each branch corresponds to three integration variables, namely the invariant mass m of the branch, its azimuthal angle ϕ, and the cosine of its polar angle θ. Every branch has its own mapping for its associated set of integration variables, and a fully assigned tree then defines a point in phase space and represents an integration channel. The module in WHIZARD implementing these trees is adequately named the `wood` phase-space module. In order to improve performance, trees with comparable phase-space mappings are grouped into groves, allowing WHIZARD to connect channels with similar characteristics when integrating over them.

For a hard scattering process, the dimension d of the phase-space manifold Π then is determined via the tree structure: Since we can only have splittings into either none or two branches, for n final-state momenta we have $(n-1)$ internal branches, each associated with a random number for each of its three associated integration variables. Because the initial first branch is special in that it has a fixed mass $m = \sqrt{s}$ associated with it, we have a dimension of

$$d = 3\,(n-1) - 1 = 3n - 4\,. \tag{3.1.3}$$

At NLO, however, we have different components with varying kinematics and therefore different phase-space dimensions (see section 3.1.1). While the Born and virtual components have n-body kinematics, the additional FKS radiation variables ξ, y, and ϕ (see eqs. (2.1.19) and (2.1.21) to (2.1.23)) of the real component each represent an additional integration variable and therefore require additional random numbers. Furthermore, the treatment of ISR and structure functions gives rise to terms that have one additional integration variable in the form of the energy fraction z of the emitted parton. For example, this is the case for the DGLAP component described in section 2.3.3. The dimensions for the different components then are [26]

$$d = 3n - 4 + \begin{cases} 0 & \text{Born, Virtual,} \\ 3 & \text{Real,} \\ 1 & \text{DGLAP.} \end{cases} \tag{3.1.4}$$

The obvious way to handle the extra integration variables, e.g. for the real component, would be to simply extend the tree structure and incorporate the additionally required mappings in the corresponding branch of the emitter. This ansatz, however, is not very well suited for the FKS phase-space mapping (see appendix A), which boosts the recoiling momenta, therefore influencing the whole tree and not only the branch the real emission is associated with. Furthermore, it would come with a significant performance overhead, as the number of trees (or phase-space channels) already present at Born level would be multiplied by the number of unique emitters.

To avoid these problems, WHIZARD applies the FKS phase-space mapping as an extension on top of the wood phase space, resulting in a direct mapping $\Phi_n \to \Phi_{n+1}$ from the n-body phase space to the $(n+1)$-body phase space for each unique emitter [26]. As we mentioned earlier in section 3.1.1, each unique emitter has its own associated term instance, containing the interaction representing the real-emission process, which will store the corresponding four-momenta obtained in this way for the $(n+1)$-kinematics.

Integration

Coming back to the general integral I_Π in eq. (3.1.1), according to the *law of large numbers*, for a large enough number of N random draws of points $x_i \in U$, with a uniform probability density on the unit hypercube U the integral I_Π can be estimated by the finite sum of its integrand such that [169]

$$\langle f \rangle_N = \frac{1}{N} \sum_{i=1}^{N} f_\varphi(x_i) \to \int_U f_\varphi(x) d^d x \,, \tag{3.1.5}$$

with

$$f_\varphi(x) \equiv f(\varphi(x)) \varphi'(x) \rho_\varphi(x) \,. \tag{3.1.6}$$

The corresponding estimate for the variance, given by

$$V_N \equiv \frac{N}{N-1} \left(\langle f^2 \rangle_N - \langle f \rangle_N^2 \right) \,, \tag{3.1.7}$$

then is equal to the square of the statistical error. During integration, WHIZARD's integrator VAMP uses a multi-channel implementation of the self-adaptive VEGAS algorithm, documented in detail in [134]. In principle, it applies a grid to the hypercube for each phase-space channel with bins along each integration dimension. Using a slightly modified version for importance sampling[6] of eq. (3.1.5), for several iterative steps of N evaluations, or *calls*, it adapts both the bin widths and the weight factors of the channel in order to reduce the variance, and thus the statistical error. This relative error can be estimated by

$$\frac{\sqrt{V_N}}{I_\Pi} = \frac{a}{\sqrt{N}} \,. \tag{3.1.8}$$

Each successful adaptation reduces the accuracy parameter a, which in turn reduces the relative error of the MC integration.

Separate and Combined Integration

There are two different modes for the NLO integration in WHIZARD: separate and combined mode. Each comes with its share of advantages and disadvantages.

The separate mode sets up an MCI for each component and computes each one separately by disabling all term instances not belonging to the current component that is to be computed. One of the advantages of this mode is that the user can specify different numbers of integration calls and iteration steps in the adaptation of the phase-space grid and event weights, and the subsequent integration runs. This can

[6] For integration, VAMP also offers a stratified sampling method [134].

lead to a much more efficient calculation. For example, on the one hand the phase space of the real emission has a higher dimension as components that have Born-like kinematics. Therefore, it needs a higher number of calls and iteration steps to reach a comparable precision. On the other hand, the virtual component might not always need as much calls and iterations, but a single call can be much slower because of the complexity the evaluation of the one-loop matrix element poses.

The separate mode also allows the user to run an instance of WHIZARD for each component, allowing them to run in parallel. This was especially relevant before the release of VAMP2 and its MPI based parallelization, however, for very complex processes one might still want to integrate the components separately since the speed-up gained by MPI begins to saturate at the order of 100 nodes.[7]

When using the separate integration mode, any subsequent event generation (see section 3.1.5) must be performed by different instances of WHIZARD for each component, as the correct normalization of the event sample cannot be guaranteed otherwise, especially if it is chosen to be dependent on the cross section.

In the combined integration mode, all term instances relevant for the full computation of the defined process are always active and share the same RNG seed for the (underlying) Born phase space. Therefore, they also share the same n-body kinematics at each point in phase space. This is in contrast to the separate mode, where the RNG seed is generally not the same between the components. Another difference is that, since the process instance sums over all active terms, the computed amplitudes of all terms are added up before being passed to the MCI. Because of this we can lose a varying amount of efficiency in the combined integration approach, since all components undergo the same number of calls and iterations, effectively meaning that the component with the slowest convergence per calls and iterations dictates the required amount of them for the whole integration.

However, the combined mode is relevant in the event generation. For example, `POWHEG` matching (see section 3.1.5) requires that all components have their contribution already evaluated for an identical phase space point of the underlying Born. This is particularly important for the Born and real contributions, as the `POWHEG` algorithm uses the ratio of both contributions to determine if the first emission in a parton shower has to be selected from the computations by the FKS subtraction or if it is handled via resummation.

A usability feature that comes with the combined integration is that for a subsequent event generation only one instance of WHIZARD is used (for reasons detailed in section 3.1.5) and the resulting event sample then is consistently normalized in this case.

[7] This is the case as of the time of this thesis. For a more explicit quantification and discussion, see [134].

3.1.3. Phase-space Setup and Singular Regions

The combinatorics of constructing the correct FKS pairs of a real flavor structure f_r with an underlying Born flavor structure f_B is of pure technical nature. We thus refrain from the technical details and only give a conceptual overview of the basic algorithm determining the divergence-inducing pairs, their associated flavor structures, and additional data. The final results are stored in the `region_data_t` type within WHIZARD. To give an example, we show its most important data structure in table 3.1 for the process $e^+e^- \to jjj$ at NLO QCD. Consulting the details of table 3.1 should help the reader to understand the details of the technical description in the following text.

alr	f_r	i_{f_r}	em	mul	$\mathcal{P}_{\text{FKS}}$	f_B	i_{f_B}
1	[11,-11,-2,2,-2, 2]	1	5	4	(3,4),(3,6),(4,5),(5,6)	[11,-11,-2,2,21]	1
2	[11,-11,-2,2,21,21]	2	3	2	(3,5),(3,6),(4,5),(4,6),(5,6)	[11,-11,-2,2,21]	1
3	[11,-11,-2,2,21,21]	2	4	2	(3,5),(3,6),(4,5),(4,6),(5,6)	[11,-11,-2,2,21]	1
4	[11,-11,-2,2,21,21]	2	5	1	(3,5),(3,6),(4,5),(4,6),(5,6)	[11,-11,-2,2,21]	1
5	[11,-11,-2,2,-1, 1]	3	5	1	(3,4),(5,6)	[11,-11,-2,2,21]	1
6	[11,-11,-1,1,-2, 2]	4	5	1	(3,4),(5,6)	[11,-11,-1,1,21]	2
7	[11,-11,-1,1,-1, 1]	5	5	4	(3,4),(3,6),(4,5),(5,6)	[11,-11,-1,1,21]	2
8	[11,-11,-1,1,21,21]	6	3	2	(3,5),(3,6),(4,5),(4,6),(5,6)	[11,-11,-1,1,21]	2
9	[11,-11,-1,1,21,21]	6	4	2	(3,5),(3,6),(4,5),(4,6),(5,6)	[11,-11,-1,1,21]	2
10	[11,-11,-1,1,21,21]	6	5	1	(3,5),(3,6),(4,5),(4,6),(5,6)	[11,-11,-1,1,21]	2

Table 3.1. The singular-region data stored in the `region_data_t` type of WHIZARD for the process $e^+e^- \to jjj$ at NLO QCD with jet-flavor content $u, \bar{u}, d, \bar{d}, g$. The meaning of the columns and the interpretation of the table as a whole is described in section 3.1.3. Particles in the flavor structures $f_{r/B}$ are represented by PDG IDs [168].

Given that we already have a list of Born flavor structures f_B in the form of PDG [168] code arrays, it is straightforward to generate all real flavor structures f_r that are possible within the selected physics model.[8] For each of the generated f_r, we check every pair of two QCD partons whether they could have originated from a splitting of another parton. If there is a valid splitting, we will store the pair of indices marking the place of the particles inside an f_r as an FKS pair (i, j). We then either replace the two partons in the real flavor structure f_r with a valid parent (in the case of $g \to q\bar{q}$) or delete the radiated gluon (in the case of $g \to gg$ and $q \to qg/\bar{q} \to \bar{q}g$). If we have a process with ISR, we might also have initial state splittings where a gluon-(anti)quark pair is replaced with the antiparticle of the (anti)quark ($g\bar{q} \leftarrow q$, see fig. 2.5). The resulting n-body flavor structure and all of its possible final-state permutations will be checked against all Born flavor structures f_B.

[8] At the moment, however, out-of-the-box NLO QCD calculations with WHIZARD only work for the SM.

Once a match is found, the FKS pair will be associated to the original real flavor structure f_r, together with further information, such as the type of splitting and whether it can induce a soft and/or collinear divergence.

We map the index i_{f_r} of the original flavor structure f_r to the corresponding index i_{f_B} representing the underlying Born flavor structure f_B. The mapping $i_{f_r} \to i_{f_B}$ is surjective such that a single fixed f_r can only ever have one underlying f_B, while several f_r might share the same underlying f_B. Thus, if we find another valid FKS pair in the same real flavor structure f_r that would lead to a different underlying Born flavor structure f_B, we store this as a new f_r with a different index i_{f_r}.

Real flavor structures f_r that differ only by permutations of final-state particles and have the same index i_{f_B} of an underlying Born flavor structure f_B associated are regarded as identical and their FKS pairs are collected in a list. Note that since an FKS pair (i, j) consists of the positional indices inside its f_r, a permutation of f_r also changes the value of i and j accordingly. If by final-state permutation one FKS pair can be transformed into another, which would lead to the same i_{f_B} before permutation, they are regarded as identical as well. However, both of them are kept in the list of FKS pairs for that real flavor structure f_r. This is important, since we compute the $\mathcal{S}$-functions (see section 2.1.2) via this list later on.

Indeed, as can be seen in the context of the trijet example in table 3.1, we also have to exchange FKS pairs between real flavor structures f_r that are associated with different underlying Born indices i_{f_B}, but are otherwise identical except for final-state permutations. When requesting a matrix element from a matrix-element generator for a (real) flavor structure, the generator will consider all possible processes leading to any final-state permutation, since it has no knowledge of the associated underlying Born flavor structure f_B. Thus, we ensure by including the FKS pairs from all possible final-state permutations of the equivalent flavor structures f_r linked to different flavor structures f_B that the $\mathcal{S}$-functions will apply the correct factors in the singular limit, as well as in non-singular phase-space regions. In the initial implementation of FKS subtraction in WHIZARD, these additional pairs were not considered, which was one of the reasons that only processes with FSR of an additional gluon gave correct results.

Finally, we permute each real flavor structure f_r such that the extra particle (or different particles as e.g for $g \to q\bar{q}/gg$) with respect to the underlying Born flavor structure f_B is(are) at the last positions and the rest of the permuted real flavor structure matches its underlying f_B. This way, an FKS pair (i, j) will have the emitting particle as index i and the radiated one as index j, as we only allow configurations where $i < j$ as defined in eq. (2.1.11). We note that for the cases where the FKS pair originated from the splitting of a third particle, present only in the underlying f_B, we will regard the lower index i of the pair (i, j) as the "emitter".

Each unique FKS pair then defines a singular α-region index `alr` that gets assigned

a multiplicity `mul`. This multiplicity represents the number of FKS pairs that belong to the real flavor structure f_r associated with a particular `alr` and that are equivalent to the FKS pair defining the `alr`. Of these equivalent pairs, the one with the highest emitter index i defines the emitter index `em` of the `alr`. Each unique `em` defines an $(n+1)$-body phase space, i.e. `alr`s with the same `em` will have the same set of kinematics.

We store the data gathered in the described way in the `region_data_t` type, shown in table 3.1 for the process $e^+e^- \to jjj$ at NLO QCD. For this example, we restrict the flavor content of the jets to u, d, g and their antiparticles, since this will produce all relevant cases for FSR in the FKS subtraction scheme for this process. We draw special attention to `alr=5,6`, where we can see the aforementioned case of having to exchange FKS pairs between two real flavor structures f_r that are identical except for final-state permutations, but have different underlying Born flavor structures f_B assigned due to the splitting associated with their respective FKS pair. More explicitly, in `alr=5`, we can see that the FKS pair `(5,6)` – associated with the emitter `em=5` of this `alr` – represents the splitting of the gluon in the underlying f_B `[11,-11,-2,2,21]` into the $\bar{d}d$-pair in the f_r `[11,-11,-2,2,-1,1]` associated with this `alr`. The FKS pair `(3,4)`, corresponding to a splitting that leads to the $\bar{u}u$-pair, would give an underlying f_B `[11,-11,-1,1,21]` and is the permuted pair `(5,6)` from `alr=6`. As stated before, this combination results in the correct application of the $\mathcal{S}$-functions to the real non-subtracted squared amplitude.

For processes with colored initial states, on the one hand we have to consider additional ISR splitting types (see section 2.2.2), and on the other hand the complexity of the region data can become huge, especially if the final state mostly consists of QCD partons. An example for the process $pp \to Zj$, with the proton and jet flavor content restricted to $u, \bar{u}, g$ for simplicity, is shown in table E.1. We already combine `alr`s where an emitted gluon could have originated from both the initial-state particles via the emitter index `em=0`. However, the number of `alr`s can still be improved, as many `alr`s produce the same amplitudes. As a future project, we implemented the experimental boolean option `?nlo_reuse_amplitudes_fks` in WHIZARD to internally summarize these equivalent `alr`s to only compute them once. Further validation, benchmarks, and improvements are still needed for this option, but first checks look promising.

3.1.4. Combinatorial Pitfalls: Symmetry Factors and Structure Functions – Two Pieces of the Same Pie

We emphasize two aspects of real counterterms: symmetry factors for identical particles in the final state and the correct application of structure functions such as PDFs. Although they are completely unrelated, the reason to mention them together comes

from the fact that for the calculation of the squared Born matrix elements in the real counterterms we refer to the underlying Born flavor structures. We do this both when requesting the squared Born matrix elements from a matrix-element generator such as OpenLoops, and when convolving the subtraction-term interaction with the structure-function interaction. The real-subtraction terms, however, describe a real flavor and momentum configuration with an unresolved parton using the squared Born matrix elements as an approximation.

Therefore, we needed to take care that WHIZARD applies the correct symmetry factor for identical final-state particles $\mathcal{N}(f_r)$ of the real flavor structure f_r (see eq. (2.2.1)) to the real-subtraction terms, since the externally provided (spin- and color-correlated) squared Born matrix elements can already contain the Born symmetry factor $\mathcal{N}(f_B)$.

Additionally, we store the Born matrix elements convolved with the structure functions corresponding to the Born configuration. In the case that we need rescaled structure functions with the flavor content of the real initial-state partons, e.g. in the case for non-soft ISR that changes the flavor of the incoming parton (see section 2.2.2 and figs. 2.5(a) and 2.5(c)), we then have to divide by the Born structure function and apply the (rescaled) one corresponding to the real flavor content. This resulting factor is calculated by the term instance and is stored in the `sf_factors` array for a quick access by the real-subtraction and the DGLAP remnant, which needs a similar treatment regarding the structure functions (see section 2.3.3).

We implemented this procedure in a way that is agnostic to the kind of structure function that is used. Therefore, it can easily be reused for the EW expansion of our FKS implementation, where electron PDFs used for calculations with massless initial-state electrons have to be treated in the same way as described above.

3.1.5. Event Generation at NLO

The term *event* usually refers to the full cascade of particle interactions and decays, starting with the hard scattering of two opposing beam constituents. The resulting particles from this hard interaction, together with particles from possible ISR, then radiate and decay further in what is called a *parton shower*, until they reach a certain energy scale below which they start to hadronize. The resulting hadrons can then decay further, until a stable state is reached. Other aspects of beam interactions can add to, or at least accompany, an event. These include multiple interactions, where more than one pair of constituents of the same two beam particles undergo an interaction, usually also referred to as the *underlying events*.

The full event information, however, is rarely obtained in a real-world experiment, as several parts are eluded due to limitations in the experimental hardware, setup, and

software. For example, detectors have to leave space for the beam axis in order to not interfere with the beam itself, leaving parts of the sphere around the interaction vertex uncovered. Additionally, different particles have different detection methods, which require their own space around the beam axis. Some particles such as neutrinos can only be inferred by their missing energy or momenta with respect to the rest of the detected particles in an event. Therefore, the challenge of the analysis of the experiment is to reconstruct the initial hard interaction by inferring the paths and types of particles in an event from their partially detected tracks and deposited energy inside of the various detectors around the beam axis. Additional uncertainties such as the exact behavior of each detector component complicate this task even further.

In a simulation via an MC event generator, however, we control each aspect of an event, and thus know exactly the momenta and other properties of every particle at every stage of the simulated event. In contrast to the experiment, which reconstructs the hard process by the decay products, an event simulation starts with the hard interaction of a process in a certain physics model. This is also called an FO calculation, as the hard interaction is described up to a certain order of the relevant coupling constants in perturbation theory (see eq. (1.0.1)). We will use the notation of fLO and fNLO whenever necessary to emphasize when we explicitly exclude the subsequent parton shower and hadronization. We also exclude the resummation of large logarithms with this notation.

Albeit they cannot fully describe experimental data without further showering and hadronization, FO differential distributions of IR-safe observables pose a physically well-defined description of perturbation theory up to the chosen order in the couplings. To obtain these distributions, MC event generators generate a number of sets of four-momenta, associated with particle flavors, spins, charges, and color flow, possible for the hard interaction in question. A single set of four-momenta and its associated data is then sometimes also referred to as an FO event. In contrast to an event in the experimental sense, it should rather be understood as a technical tool to obtain FO differential distributions.

In terms of the nomenclature established in this section and the one established in section 3.1.2, an FO event can also be interpreted as the evaluation of the function $f_\varphi(x_i)$ in eq. (3.1.6) for a single random number $x_i \in U$ from the subset U of the unit hypercube. The resulting real-valued scalar is referred to as its *weight*

$$\omega_i = f_\varphi(x_i)\,, \quad \omega_i \in \mathbb{R}\,, \quad \omega_i > 0 \text{ at LO.} \tag{3.1.9}$$

This weighting of events introduces a bias to their distribution, which is in contrast to what can be seen in a real-world experiment, where the probability distribution of events is according to the differential cross section. In an experiment, we thus observe

unweighted events, meaning that each event has a weight of $\omega_i = 1$. In order to compare simulation with experiment, the generated weighted events have to be unweighted. At LO this is straightforward, as we estimate the maximum weight $\omega_{\max}$ during the adaption of the integration. Given a random number u from a uniform distribution on the interval $(0, 1)$, we can then generate events with weight ω_i and reject those with

$$\omega_i \leq u\omega_{\max} \,. \tag{3.1.10}$$

Consequently, the remaining events all have weights equal to one, as their distribution then follows the probability distribution dictated by the squared matrix elements. Obviously, due to the rejection of events in eq. (3.1.10), generating unweighted events is not as efficient as the generation of weighted events.

At NLO, however, both the real counterterms and the virtual-subtracted squared amplitude can be negative. Therefore, the weights are not guaranteed to be positive definite, making a naive reweighting of FO events as described impossible beyond LO. This problem is addressed in the `POWHEG` matching scheme [170], for example, which was developed to realize the matching of fNLO events to a parton shower. For relatively well behaved regions of phase space, the `POWHEG` scheme can generate positive-weighted events that then are properly unweighted to resemble a physical distribution. Subsequently, those events can then be matched to a parton shower simulation by keeping the FO accuracy for the first emission, while the large logarithms in the following parton shower are resummed by modified Sudakov form factors. After showering, hadronization, decay, and detector simulation, a direct comparison between the generated and experimentally measured events is then possible.

While studies using WHIZARD at NLO with `POWHEG` matching exist [171, 172], a full implementation and validation of the fNLO QCD capabilities of WHIZARD has only been established in the course of this thesis. Since the algorithms for parton-showers, hadronization, etc. have already been realized in many external tools that can be interfaced by WHIZARD, e.g. PYTHIA, we focussed our efforts on the correct generation of FO differential distributions to build the basis upon which the implementation of `POWHEG` matching in WHIZARD can be generalized in future studies.

Fixed-order Events for Differential Distributions at NLO with WHIZARD

We explained that FO events do not fully represent actual experimental events, but rather serve as an intermediate (technical) step between the generation of the hard interaction and the subsequent parton shower. Furthermore, their exact representation at NLO strongly depends on the subtraction scheme employed by the MC event

generator and the technical details of its implementation. In addition, the demanded form of fNLO event input can differ between matching schemes and parton shower programs. Therefore, the representation of FO events by an MC event generator at NLO retains a certain ambiguity.

However, representations of FO events fulfill restrictions, like energy and momentum conservation. Furthermore, FO differential distributions of IR-safe observables computed from FO events are expected to give comparable results between different MC event generators and also with analytic calculations, because these distributions pose a physically well-defined description of perturbation theory up to the chosen order in the couplings.

In WHIZARD, the representation of fNLO events builds upon the structure of term instances presented in section 3.1.1. Depending on the selected integration method, events are either generated separately for each component or combined. For components with effectively only n-body kinematics, e.g. the Born, virtual, and DGLAP remnant, this is a straightforward task. In this case, we generate either separate independent events with n final-state momenta and associated weights dependent on the corresponding squared matrix element and kinematics, or evaluate the weight for each n-body component for the same event and associate it with the sum of the weights of the components.

For the real component, however, we have to represent the different real-emission kinematics given by the corresponding term instance for each possible unique emitter of an underlying Born flavor structure. Therefore, we have n_{em} different real-emission events, each with their own set of momenta, particle flavors, etc., and associated weight, dependent on the real, non-subtracted squared matrix element. All of them, however, share the same underlying Born kinematics, which defines their collective counterevent with a weight made up by the corresponding subtraction terms (see section 2.2). These $(n_{\text{em}} + 1)$ events – referred to as *subevents* – form an event group with the same event ID and are highly correlated. In the combined integration mode, the counterevent shares the same n-body kinematics as the Born, virtual, and DGLAP remnant component. Hence, they are combined into one single Born-like event with a weight made up from the sum of all n-body weights. This Born-like event then takes the place of the n-body counterevent in the corresponding event group.

As mentioned, we have a high correlation between subevents in an fNLO event group. In an analysis, such an event group has to be treated with care in order to ensure the correct statistical treatment when filling the weights of its subevents into binned data types such as histograms. The sum of the squared weights in one bin is used as a measure to estimate the statistical uncertainty of that bin. However, the more collinear and/or soft a real-emission event is, the larger the absolute value of its weight and the weight of its counterevent becomes. Consequently, the sum of their squared weights

would unrealistically enlarge the error estimate. Therefore, we have to add all weights of an event group that would end up in the same bin before squaring their weights and effectively count them as only one event. In this way, it is ensured that large absolute weight values due to collinear and/or soft divergences cancel out, since in these limits the real-emission event effectively has the same kinematics as the counterevent with an additional unresolved parton. Therefore, both events have a very high chance to end up in the same bin.

This does come with a caveat, however, as events and their counterevents could have kinematics very close to the edge between two bins. Numerical differences can then lead to the real-emission event landing in one bin and its counterevent in the other. The consequence of this would not only be an unrealistic enlargement of the statistical errors of both bins, but also a significant distortion of the total sum of weights of each bin. To alleviate the effects of such a misbinning, some form of weight "smearing" can be used, either already during event generation (e.g. in SHERPA [34]) or while filling bins during the analysis (e.g. in RIVET 3 [173]). Because there is no consensus yet on how to correctly implement such a smearing in WHIZARD, we have refrained from implementing it and raise the question if the handling of these misbinning effects should be left to the analysis tools.

When plotting fNLO events it is also common to encounter bins with a sum of weights below zero. Similar to the effects of misbinning, where IR divergent subevents can fall into different adjacent bins, negative bin weights are an artifact of fNLO events rather than a physical phenomenon. It can happen when predominantly Born-like subevents fall into a single bin, as both real counterevents and virtual-subtracted events can have negative weights.

An example for such a scenario would be the differential distribution of the energy of the hardest (massless) jet in the dijet process $e^+e^- \to jj$. On the one hand, counterevents have Born-like kinematics by construction and the jet energy would thus be exactly half that of the CM energy $\sqrt{s}$. On the other hand, outside of the exact soft limit, the real-emission subevents can feature a hardest jets with an energy below $\sqrt{s}/2$. Due of the fact that the sum of Born-like subevent weights at fNLO can have negative values, the sum of weights in a bin can thus become negative as well.

These negative weights are cancelled in the integration over all bins, however, and beyond fNLO adequate matching methods such as `POWHEG` take care of negative weights. At fNLO we can circumvent negative bin weights simply by choosing a suitable binning that allows for enough positive weights to counteract the negative ones.

3.2. Concluding Remarks

With the technical concepts presented in section 3.1, all essential parts needed in the understanding of how WHIZARD performs a complete fNLO QCD computation of both total cross sections and differential distributions are given. There are of course many more technical aspects of WHIZARD – such as the SINDARIN implementation, the main program loop, the LHAPDF and OLP interfaces, etc – that are important for a full fNLO computation. However, they either have already been present for LO calculations or are not essential for the general understanding of the NLO capabilities of WHIZARD. Nevertheless, during the course of the work this thesis is based upon, also significant parts of these technical aspects had to be modified to account for the parts necessary for the full fNLO QCD implementation.[9]

With the technical implementation in WHIZARD laid out, the next step is to validate its results of fNLO QCD computations, to compare it against other MC event generators, and then to apply it to phenomenological studies. This will be the topic of part II.

[9] The public Git repository for WHIZARD at [174] exhibits a full history of the changes we made.

Part II.

Numerical Studies for Fixed-NLO QCD Jet Physics

4. Validation of WHIZARD at Fixed-NLO QCD

Based upon the theoretical background of the FKS subtraction scheme, as reviewed in chapter 2, we completed the NLO QCD implementation in WHIZARD at FO. With the relevant technical details of how this has been accomplished documented in section 3.1, we now turn our attention to the validation of this implementation.

Due to the nature of WHIZARD being a *multi-purpose* MC event generator, the task of validating every possible aspect of the program for every possible process is practically impossible. Therefore, it is common practice to validate a representative and substantial subset of processes in order to increase the confidence in the implementation to a point where one can reliably assume the validity of results that have not been part of the validation themselves.

There are also many different ways to check that the program indeed works in the intended way and produces correct results. In the case of our FKS implementation, the most basic tests check that partial results of the subtraction scheme agree in certain limits or regions of phase space. A few examples for these self-consistency tests are given in section 4.1.

Comparisons against analytical results are another way to validate results, but are limited to a number of processes for which analytical calculations exist, as they quickly become unfeasible for phase spaces with higher dimensionality. However, especially for early MC computations this provided a reliable way to verify results.

Of course, beyond FO, MC simulations have to be validated against actual experimental data. At NLO QCD, most major multi-purpose MC event generators have undergone extensive validation in this regard over the past decades, with a multitude of data available from e.g. LEP or LHC runs. Unfortunately, our NLO implementation in WHIZARD so far only supports FO calculations. While it would in principle be possible to some degree to unfold the experimental data onto the hard process, this would be impractical and introduce additional possibilities for errors. However, since there exist extensively-validated multi-purpose MC event generators that are capable of FO calculations, we can validate our results against theirs for a multitude of processes.

In section 4.2 we compare fNLO QCD cross sections computed by WHIZARD against results from the multi-purpose MC event generator MG5_aMC@NLO [33]

for a vast selection of lepton and hadron-collider processes. This is followed by an exhaustive comparison of differential distributions for exclusive numbers of jets in multi-jet production at a future 1 TeV lepton collider in section 4.3. There, we additionally compare against the multi-purpose MC event generator SHERPA [34, 175]. We also briefly consider differential distributions for hadron-collider processes by means of the process $pp \to Zj$ in section 4.3.5.

It is important to note that while the comparisons of results of WHIZARD to those of MG5_aMC@NLO and SHERPA in section 4.2 and section 4.3 primarily serve the purpose of validating our FKS implementation in WHIZARD, they also provide a much more general cross-check of the physics and technical concepts implemented in each program considered in this comparison. As they become more and more complex over time, with significantly different approaches to physical and technical problems, it becomes increasingly important to check, understand, and keep track of these differences across MC event generators. This is important, as these programs are frequently used in state-of-the-art studies, both in theory and at experiments, where conclusions and discoveries can very well depend on the correct physics description in any part of an MC program. As an example, a recent study [176] that compared several well-known MC event generators in the context of vector-boson scattering at the LHC illustrated the need to understand differences across these tools in order to make meaningful theoretical predictions.

4.1. Self-consistency Tests

Before comparing full results for fNLO total cross sections and differential distributions to other MC event generators, we can conduct self-consistency tests to check if the FKS implementation performs in a technically expected way. In the following we will present the three classes of self-consistency tests that we predominantly used to check the technical implementation of the FKS subtraction scheme in WHIZARD. While there are certainly many other possible ways to check that the implementation works as expected, we found the most telling ones to be these presented here.

4.1.1. Limit Checks

The first test from which we can deduce information about the correctness of our FKS implementation is a check of the cancellation of the real non-subtracted squared amplitude $\hat{\mathcal{R}}_\alpha(\xi, y)$ and the corresponding counterterm (see eq. (2.2.5)), both evaluated for a kinematic configuration approaching either a soft, collinear, or soft-collinear limit. In these limits, the additional parton is unresolvable and thus the real subtracted contribution has to approach zero. If this were not the case and the difference in

these limits between $\hat{\mathcal{R}}_\alpha(\xi, y)$ and the corresponding counterterm becomes larger than numerical fluctuations, the integration would not converge. We give an example for a good convergence in listing F.2, where we show the integration of the real component for the process $e^+e^+ \to jjj$.

In WHIZARD, the limit test as described can be conducted for the real component via the command-line debug flag `--debug subtraction` and the SINDARIN commands

- `?test_coll_limit` (setting $\xi = 0.5, y = 1 - 10^{-7}$),
- `?test_anti_coll_limit` (setting $\xi = 0.5, y = -(1 - 10^{-7})$),
- and `?test_soft_limit` (setting $\xi = 10^{-5}, y = 0.5$),

which can be set to `true` or `false` individually. By setting the soft test case to `true` together with either the collinear or anti-collinear test case, we can test the soft-collinear limit where $\xi = 10^{-5}$ and $y = \pm(1 - 10^{-7})$. Because of possible numerical instabilities in the exact limits, we do not set $\xi = 0$ or $y = \pm 1$ for the real non-subtracted squared amplitude. WHIZARD then computes $\hat{\mathcal{R}}_\alpha(\xi, y)$ and the respective counterterm $\hat{\mathcal{R}}^{\text{cnt}}_\alpha(\xi, y)$ for the selected limit and compares them via

$$\frac{\left|\hat{\mathcal{R}}_\alpha(\xi, y) - \hat{\mathcal{R}}^{\text{cnt}}_\alpha(\xi, y)\right|}{\max\left(\left|\hat{\mathcal{R}}_\alpha(\xi, y)\right|, \left|\hat{\mathcal{R}}^{\text{cnt}}_\alpha(\xi, y)\right|\right)} \overset{!}{<} 0.01\,. \tag{4.1.1}$$

This acceptance level is motivated only by empirical observation and experience from the numerics of a wide variety of NLO QCD processes and their IR limits. Also, the relative number of calls passing these acceptance levels can vary significantly, as processes with e.g. more complex phase-space configurations might need an absolute value of y closer to one or a value of ξ closer to zero to pass. For most processes, however, we observed that if roughly $80 - 100\%$ of calls for a specific singular α-region `alr` (see section 3.1.3) in each limit fulfill these acceptance levels with the chosen values for ξ and y, the integration shows an excellent convergence.

4.1.2. Closure Tests

While the limit checks cover the divergent regions of phase space and thus can be indicative of the quality of convergence in the integration, they cannot give clues on any possible misbehavior of the FKS implementation outside of these extreme regions of phase space. For example, as we have discussed in section 3.1.3, an FKS pair that does not belong to the emitter of an α-region `alr` will give a trivial contribution to the calculation of the $\mathcal{S}$-function in a singular region of phase space, while everywhere else it gives a non-zero contribution. If we had missed implementing this pair, the limit

checks would still work fine, as would the convergence of the integration. The final result, however, would not be correct.

In order to catch these kinds of errors, at least for the real component, we can conduct what we call *closure tests*. Basically, what this means is that the integrated cross section for a real emission process should be equal to the cross section we would get if we would treat this process as Born-level, given that we impose the same cuts and jet selections such that the Born integral is finite.

For an explicit example let us look at the real-emission process as a correction to the process $e^+e^- \to jjj$ at fNLO QCD and the Born process of $e^+e^- \to jjjj$, where we excluded top quarks from the jets j. Also note that at FO, the jets are obtained simply from parton momenta onto which we apply an IR-safe jet algorithm. Both processes are of the order $\mathcal{O}(\alpha^2\alpha_S^2)$ in the QED and strong coupling constants, α and α_S, respectively. We refer to section 4.2.1 for the exact setup of these processes, as the details are not relevant at this point. We note, however, that we have the same setup for both processes and require four distinctive – i.e. resolved – jets in both cases. In this way, we compute the real emission of $e^+e^- \to jjj$ only in the phase-space continuum, while cutting all singular configurations where a parton becomes sufficiently soft and/or collinear enough to another parton that they are clustered into one jet object. For a CM energy of 1 TeV, the results are

$$\begin{aligned} \sigma^{\text{real}}_{e^+e^-\to jjj} &= 104.440(26)\ \text{fb}, \\ \sigma^{\text{Born}}_{e^+e^-\to jjjj} &= 104.456(23)\ \text{fb}, \end{aligned} \tag{4.1.2}$$

and agree perfectly within their statistical errors. As another example, in the same manner we also compare the cross section for the real emission of $pp \to Zj$ with the Born cross section of $pp \to Zjj$. In this way, we also probe the ISR part of the FKS implementation. Again, the results agree perfectly within their statistical errors as shown for a CM energy of 13 TeV here:

$$\begin{aligned} \sigma^{\text{real}}_{pp\to Zj} &= 2.3611(28) \cdot 10^6\ \text{fb}, \\ \sigma^{\text{Born}}_{pp\to Zjj} &= 2.3623(20) \cdot 10^6\ \text{fb}. \end{aligned} \tag{4.1.3}$$

4.1.3. Check for Independence of Free Parameters

In the FKS subtraction scheme, several free parameters enter the cross section of the hard interaction. These originate from various ambiguities introduced in the derivation of the analytical FKS terms. As one example, because of the ambiguity in the partition of unity (see eq. (2.1.25)), the $\mathcal{S}$-functions depend on an arbitrary exponent p in eq. (2.1.37). Another example is that the definitions of the modified plus distributions in eqs. (2.2.12) to (2.2.14) include the arbitrary parameters ξ_{cut}, as well as δ_I and δ_O,

that enter the different terms in eq. (2.2.17). These parameters effectively act as a phase-space slicing between the singular and non-singular regions of the real-emission phase space. While the choice of the values of these free parameters can influence the value for specific terms of the FKS subtraction scheme, the total cross section should be independent of this choice. Thus, we are provided with a perfect opportunity to test the FKS implementation by varying these parameters.

As an example, for the process $e^+e^- \to jjjj$ we test the variation of δ_O, which effectively produces a shift between final-state contributions from the real collinear and soft-collinear counterterms (see eq. (2.2.20)), and the integrated collinear counterterms (see section 2.3.2). For the same process setup that will be detailed in section 4.2.1, we compute the real and virtual contributions for several choices of δ_O, shown in table 4.1. As expected, we can observe that within the statistical errors of the MC computation their sum remains independent of the choice of δ_O.

δ_O	σ^{real}[fb]	σ^{virt}[fb]	$\sigma^{\text{real+virt}}$[fb]
2.00	57.24(16)	−113.11(12)	−55.87(20)
1.50	32.19(17)	−88.30(10)	−56.11(20)
0.70	−32.84(18)	−22.95(5)	−55.79(18)
0.25	−121.09(18)	65.32(6)	−55.77(19)
0.05	−259.36(22)	203.37(18)	−55.99(29)

Table 4.1. Cross sections for the real and virtual contributions to the process $e^+e^- \to jjjj$ at NLO QCD for variations of δ_O. The sum of their contributions is shown as well. See section 4.1.3 for details.

4.2. Comparison of Fixed-order Cross Sections

The obvious next step in a thorough validation of the WHIZARD fNLO QCD capabilities is to cross-check computed cross sections with the results from other MC event generators. Conveniently, there exists an extensive compilation of fNLO QCD total cross section results, conducted by the MadGraph developers for the validation of their implementation of FKS in the MG5_aMC@NLO program [33] and documented in [33, tables 1 to 11]. In there, scattering processes with both leptonic and hadronic initial states were considered.

In the following, we will first detail the process and parameter setup for the FO cross section comparison between WHIZARD and MG5_aMC@NLO in section 4.2.1, and subsequently present and discuss the results in section 4.2.2. It has to be noted that we could have used e.g. Sherpa as a third MC event generator to give an additional cross-check between different subtraction schemes, as Sherpa uses the CS

subtraction scheme. However, the results in [33] were calculated for a dynamical scale. Because SHERPA implements a different running of α_S, as discussed in more detail in section 4.3.1, any comparison including SHERPA would have required additional scale variations for each process to assess the agreement between the programs. This proved to be prohibitive within the time scope of this thesis due to the large number of processes considered in this comparison of cross sections.

For a technical validation of fNLO QCD processes, a restriction to a physically limited setup is very beneficial. Hence, building upon the setup in [33], we limit ourselves to the hard fNLO QCD process without applications of EW contributions, parton showers or hadronization.

4.2.1. Process and Parameter Setup

In order to properly compare the WHIZARD results with MG5_aMC@NLO, we translated all of the process setup as described in [33] into SINDARIN code, taking into account further details given in the MADGRAPH run and parameter cards [177] used in that paper. In the following, we will discuss the details of the process and parameter setup.

Most of the physical parameter setup can directly be read from [177] and transferred to a SINDARIN file. Lepton-collider processes are computed for a CM energy of $\sqrt{s} = 1\,\text{TeV}$, whereas results for hadron-collider processes are presented for $\sqrt{s} = 13\,\text{TeV}$. If not stated otherwise, quarks are considered massless, with the exception of the top quark. Thus, we default to the five-flavor scheme as described in [178, 179] and references therein. Resonant-top contributions in hadron-collider processes with at least two W bosons are avoided by calculating them in the four-flavor scheme. The top mass, and the Higgs and EW gauge boson masses are set as

$$\begin{aligned} m_H &= 125\ \text{GeV}, & m_t &= 173.2\ \text{GeV}, \\ m_W &= 80.419002\ \text{GeV}, & m_Z &= 91.188\ \text{GeV}. \end{aligned} \tag{4.2.1}$$

Leptons are considered massless, with the exception of the tau mass being set to

$$m_\tau = 1.777\ \text{GeV}. \tag{4.2.2}$$

Although both MG5_aMC@NLO and WHIZARD have implementations to handle intermediate resonances via the complex mass scheme, top quarks, as well as the EW bosons W and Z are considered to be stable and their widths are consequently set to zero for this comparison.

The EW scheme is chosen such that m_W, m_Z, and

$$G_F = 1.16639 \cdot 10^{-5} \tag{4.2.3}$$

are the input parameters. Furthermore, we assume a diagonal CKM matrix.

For the central scale $\mu = \mu_R = \mu_F$, we compute the scalar sum H_T of the transverse masses such that

$$\mu = \frac{1}{2} H_T \,, \quad H_T \equiv \sum_i \sqrt{p_{\mathrm{T},i}^2 + m_i^2} \,, \tag{4.2.4}$$

where the index i runs over all particles in the final state – including those that are not strongly interacting. For the FO cross section comparison between MC event generators in this work, we do not consider scale variations. Any errors given are purely the statistical MC errors reported by each program.

Both fLO and fNLO hadron collision results have been computed with the central PDF of the set

- MSTWnlo2008 with errors at 68% confidence level (CL) [180],

with the value of the strong coupling $\alpha_S(m_Z)$ being set by the PDF. For lepton collisions we set

$$\alpha_S(m_Z) = 0.118 \,. \tag{4.2.5}$$

The internal running of α_S in WHIZARD is calculated up to $\mathcal{O}(\alpha_S^2)$, which to our knowledge equals the internal settings of MG5_aMC@NLO.

As already stated in detail in the beginning of chapter 2, care has to be taken in regard to the selection of EW and strong coupling orders. Otherwise, a consistent NLO QCD calculation is not guaranteed, as gauge invariance might be violated. When running MG5_aMC@NLO at fNLO QCD, the program automatically only selects the Feynman diagrams with the highest order possible in the strong coupling constant α_S for a process in question, thus guaranteeing that the real emission leading to an additional factor in the strong coupling is a pure QCD correction (see fig. 2.1).

In WHIZARD, we have to explicitly set the coupling orders for the Born process via Sindarin commands. At the time of writing this thesis, this is only possible for external matrix-element generators, but not for O'Mega. Also, as stated in the introduction of chapter 3, the fNLO QCD implementation in WHIZARD so far is only fully complete for OpenLoops 2 as the matrix-element generator, to which we will default for all components in WHIZARD throughout this work (see section 3.1.1).

For the finite one-loop contributions, MG5_aMC@NLO comes with its own one-loop provider MadLoop [49].

The anti-k_T jet-clustering algorithm [181] is imposed on all massless final state partons, where the jets j are selected by the restrictions

$$R = 0.5\,, \quad p_{\mathrm{T}}(j) > 30\ \mathrm{GeV}, \quad |\eta(j)| < 4\,, \tag{4.2.6}$$

with the jet-radius (or jet-resolution) parameter R. If the number of jets fulfilling these conditions are less than the number of jets in the Born process, we reject the event. For the 13 TeV hadron-collider processes with pure multi-jet final states we impose stricter cuts, with as many jets as are present at Born level fulfilling $p_{\mathrm{T}}(j) > 80\,\mathrm{GeV}$, out of which at least one jet has to have $p_{\mathrm{T}}(j) > 100\,\mathrm{GeV}$.

It is important to note that for the real component we cut real-emission and Born-like phase-space points separately, which is the default behavior of both MG5_aMC@NLO and WHIZARD. This is necessary, as there exist cases where the real emission does not pass the jet selection, while the n-body kinematics of the underlying Born passes. If we would cut the Born-like phase-space point as well, the cancellation between the integrated subtraction terms and the real counterterms would not resemble an addition of zero anymore, leading to results that are incorrectly dependent on the underlying subtraction mechanism.

The results in [33, tables 1 to 11] also contain processes with bottom quarks and photons in the final state. The former have been computed with a specific selection on the number of b-jets, while the latter have been done with Frixione isolation [182, 183] in order to isolate the photons from jets present in the final state. Although these two selection criteria have been implemented in WHIZARD recently, they have not been validated yet. However, they are not an NLO-specific feature and thus are not related to the validity of our FKS implementation.

Also, no (complete) OpenLoops libraries for the processes $pp \to W^+W^-W^+W^-(4f)$, $pp \to ZZZZ$, and $pp \to t\bar{t}W^{\pm}Z$ were available. These processes have been excluded from the comparison as well.

Finally, in [33] fNLO QCD cross sections for hadron-collider processes with both Higgs boson and single-top quark production are presented as well. Those have been left out in our comparison so far – with the exception of $pp \to HZ$ – simply due to time constraints and the fact that we do not anticipate any unexpected results caused by our FKS implementation.

4.2.2. Comparison and Discussion of Results

The results for the FO cross sections, computed as described in section 4.2.1, are presented in this section. Unless we explicitly state otherwise, the LO results for MG5_aMC@NLO are taken directly from [33]. For the NLO results, however, the

version of MG5_aMC@NLO used to compute the results in [33] (and every subsequent version up to 2.6.2) had a bug present for processes having identical QCD particles in the final state at Born level [184], leading to significant deviations from the actual result. We thus recalculated every MG5_aMC@NLO NLO cross section with version 2.7.3.[1] The only exception had to be made for the process $pp \to W^{\pm}jjj$, which could not be recalculated because of unresolved technical difficulties with MG5_aMC@NLO at the time of this thesis and we resorted to the result from the paper for this process.

Roughly following the ordering of processes in [33], we grouped together processes with similar characteristics in the presentation of the results. In table 4.2 we state the results for lepton-collider processes with pure QCD final states, consisting of heavy top quarks and light jets. Table 4.3 presents results for lepton-collider processes with a top quark pair and one heavy scalar or up to two vector bosons in the final state, accompanied by up to two light jets or an additional boson. Hadron-collider results are next, with up to two massive vector bosons in the final state, and up to two additional light jets, given in table 4.4. Final states with three heavy vector bosons and up to one light jet, as well as combinations of four final-state heavy vector bosons, are collected in table 4.5. Finally, in table 4.6 we conclude the hadron-collider processes with pure QCD final states, and top quark pairs with up to one light jet or an additional heavy vector boson. The $pp \to HZ$ result is given there as well.

The results for MG5_aMC@NLO and WHIZARD are presented with the K-factor, defined as

$$K \equiv \frac{\sigma_{\mathrm{NLO}}}{\sigma_{\mathrm{LO}}}\,, \tag{4.2.7}$$

as well as the significance $\sigma^{\mathrm{sig}}_{\mathrm{NLO}}$ of the deviation between two results, given by

$$\sigma^{\mathrm{sig}}_{\mathrm{NLO}} \equiv \frac{\left|\sigma^{\mathrm{MG}}_{\mathrm{NLO}} - \sigma^{\mathrm{WH}}_{\mathrm{NLO}}\right|}{\sqrt{\left(\Delta\sigma^{\mathrm{MG}}_{\mathrm{NLO}}\right)^2 + \left(\Delta\sigma^{\mathrm{WH}}_{\mathrm{NLO}}\right)^2}}\,, \tag{4.2.8}$$

with the total fNLO QCD cross sections $\sigma^{\mathrm{MG/WH}}_{\mathrm{NLO}}$ and their statistical errors $\Delta\sigma^{\mathrm{MG/WH}}_{\mathrm{NLO}}$ of MG5_aMC@NLO and WHIZARD, respectively. It is common to denote the significance as $r\sigma^{\mathrm{sig}}_{\mathrm{NLO}}$, with r being the relative factor to the case where $\sigma^{\mathrm{sig}}_{\mathrm{NLO}} = 1$. Values above $3\sigma^{\mathrm{sig}}_{\mathrm{NLO}}$ can in general be interpreted as a hint for a deviation, with the commonly accepted threshold for a statistically significant deviation being $5\sigma^{\mathrm{sig}}_{\mathrm{NLO}}$.

In the following, we will discuss the results in more detail. Considering lepton-collider processes in table 4.2 first, we can exclusively test the FSR part of our FKS implementation. The processes with pure light-jet final states provide an excellent

[1] Some results have been obtained with versions between 2.6.2 and 2.7.3, as the program has received several updates during the time of this thesis. However, we do not see any deviations between those versions themselves.

test case for massless FKS subtraction. Starting with the dijet process, we have a very simple phase space and the only possible QCD radiation is the emission of a soft and/or collinear gluon. These splittings do not introduce spin correlations and color correlations are minimal. Adding one jet at Born level introduces $g \to q\bar{q}/gg$ splittings at NLO and thus spin correlations.

Increasing the number of QCD jets also leads to more complicated color correlations. Furthermore, $e^+e^- \to jjj$ (or trijet) is the first process where the treatment of FKS pairs becomes non-trivial, as explained in section 3.1.3 regarding `alr=5,6` in table 3.1. Any wrong treatment of identical particles in the real-emission final state would become noticeable here as well, due to the $g \to gg$ splitting. While increasing the number of jets in the Born process does not introduce new aspects of FKS, the complexity of the aforementioned features rises significantly. Therefore, we would expect even the smallest discrepancy from the correct implementation to be reflected in an obvious disagreement between MG5_aMC@NLO and WHIZARD. The fact that we do not observe any such deviation even for high jet multiplicities already gives high confidence in our implementation.

The correct treatment of massive QCD partons is tested with the inclusion of top-quark pairs in the final state. Without any additional light jets, we do not have spin correlations. In addition, the mass of the top quark regularizes any collinear divergences. Therefore, the only divergences originate from the emission of a soft gluon. Moreover, these processes test the correct implementation of the massive-massive eikonal integrals in the virtual-subtracted squared amplitude (see section 2.3.1 and appendix D.2).

The implementation of mixed final states with both massive and massless QCD partons is tested by the addition of up to three light jets to processes with top quark pairs. These processes provide validation of out implementation of the massive-massless eikonal integrals.

More complex phase-space configurations arise when the emitter-emitted system recoils against additional heavy final-state bosons in processes as shown in table 4.3. Although these processes have rather simple FKS combinatorics regarding the singular regions, they can have complicated loop diagrams. Therefore, the observed excellent agreement with MG5_aMC@NLO can additionally be seen as a validation of the correct interplay between WHIZARD and OpenLoops.

QCD corrections involving ISR are scrutinized by calculating total cross sections for hadron-collider processes, with the results given in tables 4.4 to 4.6. Pure ISR is produced by considering colorless final states at Born level, e.g. processes with a mixture of up to four heavy scalar and/or vector bosons in the final state, with the simplest being $pp \to Z$. Higher boson multiplicity again does not add a higher complexity in FKS combinatorics, but rather tests the integrity of our phase-space implementation and the handling of complex (loop) diagrams (via the OLP interface).

Process	MG5_aMC@NLO			WHIZARD			
Top quarks plus jets	σ_{LO}[fb]	σ_{NLO}[fb]	K	σ_{LO}[fb]	σ_{NLO}[fb]	K	$\sigma^{\mathrm{sig}}_{\mathrm{NLO}}$
$e^+e^- \to jj$	622.70(5)	639.30(12)	1.03	622.737(8)	639.39(5)	1.03	0.69
$e^+e^- \to jjj$	340.4(7)	317.3(8)	0.93	340.6(5)	317.8(5)	0.93	0.53
$e^+e^- \to jjjj$	104.09(20)	103.67(26)	1.00	105.0(3)	104.2(4)	0.99	1.11
$e^+e^- \to jjjjj$	22.35(5)	24.65(4)	1.10	22.33(5)	24.57(7)	1.10	0.99
$e^+e^- \to t\bar{t}$	166.32(11)	174.5(3)	1.05	166.37(12)	174.55(20)	1.05	0.14
$e^+e^- \to t\bar{t}j$	47.95(9)	53.336(10)	1.11	48.12(5)	53.41(7)	1.11	1.05
$e^+e^- \to t\bar{t}jj$	8.608(18)	10.515(19)	1.22	8.592(19)	10.526(21)	1.23	0.39
$e^+e^- \to t\bar{t}jjj$	1.0371(21)	1.415(4)	1.36	1.035(4)	1.405(5)	1.36	1.56
$e^+e^- \to t\bar{t}t\bar{t}$	$0.6385(12) \cdot 10^{-3}$	$1.1941(20) \cdot 10^{-3}$	1.87	$0.6388(8) \cdot 10^{-3}$	$1.1922(11) \cdot 10^{-3}$	1.87	0.83
$e^+e^- \to t\bar{t}t\bar{t}j$	$2.662(4) \cdot 10^{-5}$	$5.264(9) \cdot 10^{-5}$	1.98	$2.673(7) \cdot 10^{-5}$	$5.251(11) \cdot 10^{-5}$	1.96	0.91

Table 4.2. Comparison between MG5_aMC@NLO and WHIZARD of total FO cross sections at LO and NLO QCD for light-jet production, and top quark production in association with light jets. Results are shown for a 1 TeV e^+e^- collider, with cuts applied as explained in section 4.2.1. Statisical errors of the integration, as well as the individual K-factor for each program, are shown. For better comparison, we provide the significance $\sigma^{\mathrm{sig}}_{\mathrm{NLO}}$ of the deviation between the NLO results in the last column. Both LO and NLO MG5_aMC@NLO results are recalculated as mentioned in section 4.2.2.

Process	MG5_aMC@NLO			WHIZARD			
Top quarks plus bosons	σ_{LO}[fb]	σ_{NLO}[fb]	K	σ_{LO}[fb]	σ_{NLO}[fb]	K	$\sigma^{\mathrm{sig}}_{\mathrm{NLO}}$
$e^+e^- \to t\bar{t}H$	2.021(4)	1.909(3)	0.94	2.020(3)	1.912(3)	0.95	0.71
$e^+e^- \to t\bar{t}Hj$	$2.545(5) \cdot 10^{-1}$	$2.665(6) \cdot 10^{-1}$	1.05	$2.536(4) \cdot 10^{-1}$	$2.657(4) \cdot 10^{-1}$	1.05	1.11
$e^+e^- \to t\bar{t}Hjj$	$2.665(5) \cdot 10^{-2}$	$3.141(9) \cdot 10^{-2}$	1.18	$2.646(8) \cdot 10^{-2}$	$3.123(9) \cdot 10^{-2}$	1.18	1.41
$e^+e^- \to t\bar{t}Z$	4.630(8)	4.942(11)	1.07	4.638(3)	4.937(3)	1.06	0.44
$e^+e^- \to t\bar{t}Zj$	$6.043(10) \cdot 10^{-1}$	$6.917(24) \cdot 10^{-1}$	1.14	$6.027(9) \cdot 10^{-1}$	$6.921(11) \cdot 10^{-1}$	1.15	0.15
$e^+e^- \to t\bar{t}Zjj$	$6.426(16) \cdot 10^{-2}$	$8.181(21) \cdot 10^{-2}$	1.27	$6.436(21) \cdot 10^{-2}$	$8.241(29) \cdot 10^{-2}$	1.28	1.68
$e^+e^- \to t\bar{t}W^{\pm}jj$	$2.372(5) \cdot 10^{-4}$	$3.714(8) \cdot 10^{-4}$	1.57	$2.387(8) \cdot 10^{-4}$	$3.716(10) \cdot 10^{-4}$	1.56	0.16
$e^+e^- \to t\bar{t}HZ$	$3.611(7) \cdot 10^{-2}$	$3.592(7) \cdot 10^{-2}$	0.99	$3.623(19) \cdot 10^{-2}$	$3.584(19) \cdot 10^{-2}$	0.99	0.40
$e^+e^- \to t\bar{t}ZZ$	$3.774(7) \cdot 10^{-2}$	$4.027(8) \cdot 10^{-2}$	1.07	$3.788(6) \cdot 10^{-2}$	$4.032(7) \cdot 10^{-2}$	1.06	0.47
$e^+e^- \to t\bar{t}HH$	$1.3635(25) \cdot 10^{-2}$	$1.2140(24) \cdot 10^{-2}$	0.89	$1.3650(15) \cdot 10^{-2}$	$1.2168(16) \cdot 10^{-2}$	0.89	0.97
$e^+e^- \to t\bar{t}W^+W^-$	$1.3632(22) \cdot 10^{-1}$	$1.5350(25) \cdot 10^{-1}$	1.13	$1.3672(21) \cdot 10^{-1}$	$1.5385(22) \cdot 10^{-1}$	1.13	1.05

Table 4.3. Comparison between MG5_aMC@NLO and WHIZARD of total FO cross sections at LO and NLO QCD for top-quark production in association with heavy bosons. Results are shown for a 1 TeV e^+e^- collider, with cuts applied as explained in section 4.2.1. Statisical errors of the integration, as well as the individual K-factor for each program, are shown. For better comparison, we provide the significance $\sigma^{\mathrm{sig}}_{\mathrm{NLO}}$ of the deviation between the NLO results in the last column. Both LO and NLO MG5_aMC@NLO results are recalculated as mentioned in section 4.2.2.

The correct treatment of different quark flavors in the initial state is validated by processes with at least one final-state $W^{\pm}$ boson.

We observe an excellent agreement between the MG5_aMC@NLO and WHIZARD results, with the only outlier being a deviation with a significance of $3.56\sigma^{\text{sig}}_{\text{NLO}}$ for the process $pp \to W^+W^-ZZ(4f)$ as presented in table 4.5. The question why we see a disagreement here (and only here) is not fully answered yet and is part of an ongoing discussion. We recomputed both results several times with varying seeds to exclude statistical fluctuations. This, however, did not reduce $\sigma^{\text{sig}}_{\text{NLO}}$. It is possible that for this process in particular the different phase-space techniques of both MC programs disagree enough to introduce a slight deviation that is reflected in the total cross section. Other possibilities include differences in the EW scheme or a problem with the OpenLoops library for this particular process. Last, but not least, there always exists the possibility of an error in this specific process setup we simply did not catch.

In relation to the vast amount of processes where we agree with MG5_aMC@NLO, however, this outlier remains an anomaly that is significant enough to warrant a second look, but is insignificant enough to not invalidate our FKS implementation as a whole.

The addition of up to two light jets to the pure bosonic final states in tables 4.4 and 4.5 leads to processes that provide a perfect test case for the interplay of initial- and final-state QCD corrections. Again, the heavy bosons simply act as recoilers to the emitter-emitted system, giving rise to more complex kinematics and loop diagrams.

The highest amount of complexity regarding the combinatorics in the FKS subtraction scheme, however, can be found in processes like $pp \to jj$ and $pp \to jjj$, listed in table 4.6. There, the number of singular regions easily reaches four digits. The fact that we do not see any significant deviation from the MG5_aMC@NLO results here as well is a strong confirmation of our correct implementation of massless FKS subtraction throughout all components in WHIZARD.

Finally, massive and massive-massless QCD final states in the context of hadron collisions are validated via processes containing top-quark pairs in the final state as shown in table 4.6, with the addition of massive bosons and/or light jets. We observe a perfect agreement with MG5_aMC@NLO.

In summary, for almost any process in this comparison, we observe an excellent agreement between MG5_aMC@NLO and WHIZARD both at LO and NLO. Only a very few processes deviate by slightly more than $2\sigma^{\text{sig}}_{\text{NLO}}$. Statistically, this is expected for the number of processes considered in this comparison. We conclude that our implementation of FKS subtraction correctly computes total cross sections at fNLO QCD.

Process	MG5_aMC@NLO			WHIZARD			
Vector boson (pair) plus jets	σ_{LO}[fb]	σ_{NLO}[fb]	K	σ_{LO}[fb]	σ_{NLO}[fb]	K	$\sigma^{\mathrm{sig}}_{\mathrm{NLO}}$
$pp \to W^{\pm}$ *	$1.370(3) \cdot 10^8$	$1.765(5) \cdot 10^8$	1.29	$1.3749(8) \cdot 10^8$	$1.7696(10) \cdot 10^8$	1.29	0.90
$pp \to W^{\pm}j$ *	$2.045(4) \cdot 10^7$	$2.839(9) \cdot 10^7$	1.39	$2.046(3) \cdot 10^7$	$2.854(5) \cdot 10^7$	1.39	1.46
$pp \to W^{\pm}jj$	$6.805(15) \cdot 10^6$	$7.780(13) \cdot 10^6$	1.14	$6.856(12) \cdot 10^6$	$7.814(27) \cdot 10^6$	1.14	1.13
$pp \to W^{\pm}jjj$ †	$1.821(2) \cdot 10^6$	$2.005(8) \cdot 10^6$	1.10	$1.840(5) \cdot 10^6$	$1.978(7) \cdot 10^6$	1.07	2.54
$pp \to Z$	$4.248(5) \cdot 10^7$	$5.415(16) \cdot 10^7$	1.27	$4.2541(3) \cdot 10^7$	$5.4086(16) \cdot 10^7$	1.27	0.40
$pp \to Zj$	$7.209(5) \cdot 10^6$	$9.75(3) \cdot 10^6$	1.35	$7.215(4) \cdot 10^6$	$9.733(10) \cdot 10^6$	1.35	0.54
$pp \to Zjj$	$2.348(6) \cdot 10^6$	$2.684(5) \cdot 10^6$	1.14	$2.364(5) \cdot 10^6$	$2.676(7) \cdot 10^6$	1.13	0.93
$pp \to Zjjj$	$6.344(8) \cdot 10^5$	$6.897(22) \cdot 10^5$	1.09	$6.381(23) \cdot 10^5$	$6.85(3) \cdot 10^5$	1.07	1.26
$pp \to W^+W^-(4f)$	$7.355(5) \cdot 10^4$	$10.296(28) \cdot 10^4$	1.40	$7.352(10) \cdot 10^4$	$10.268(11) \cdot 10^4$	1.40	0.93
$pp \to W^+W^-j(4f)$	$2.865(3) \cdot 10^4$	$3.716(10) \cdot 10^4$	1.30	$2.853(7) \cdot 10^4$	$3.733(7) \cdot 10^4$	1.31	1.39
$pp \to W^+W^-jj(4f)$ *	$1.1480(24) \cdot 10^4$	$1.3866(24) \cdot 10^4$	1.21	$1.150(5) \cdot 10^4$	$1.372(6) \cdot 10^4$	1.19	2.26
$pp \to W^+W^+jj$ *	$1.5011(19) \cdot 10^2$	$2.252(4) \cdot 10^2$	1.50	$1.506(5) \cdot 10^2$	$2.235(7) \cdot 10^2$	1.48	2.11
$pp \to W^-W^-jj$	$6.752(7) \cdot 10^1$	$9.994(14) \cdot 10^1$	1.48	$6.772(24) \cdot 10^1$	$9.982(28) \cdot 10^1$	1.47	0.38
$pp \to ZW^{\pm}$	$2.777(3) \cdot 10^4$	$4.485(12) \cdot 10^4$	1.62	$2.780(5) \cdot 10^4$	$4.488(4) \cdot 10^4$	1.61	0.24
$pp \to ZW^{\pm}j$	$1.605(5) \cdot 10^4$	$2.100(5) \cdot 10^4$	1.31	$1.609(4) \cdot 10^4$	$2.0940(28) \cdot 10^4$	1.30	1.05
$pp \to ZW^{\pm}jj$	$8.038(9) \cdot 10^3$	$9.05(2) \cdot 10^3$	1.13	$8.06(3) \cdot 10^3$	$9.02(4) \cdot 10^3$	1.12	0.67
$pp \to ZZ$ *	$1.0971(20) \cdot 10^4$	$1.4185(25) \cdot 10^4$	1.29	$1.0969(10) \cdot 10^4$	$1.4183(11) \cdot 10^4$	1.29	0.07
$pp \to ZZj$	$3.662(3) \cdot 10^3$	$4.817(16) \cdot 10^3$	1.32	$3.667(9) \cdot 10^3$	$4.807(8) \cdot 10^3$	1.31	0.56
$pp \to ZZjj$ *	$1.3454(24) \cdot 10^3$	$1.678(3) \cdot 10^3$	1.25	$1.356(6) \cdot 10^3$	$1.684(8) \cdot 10^3$	1.24	0.70

Table 4.4. Comparison between MG5_aMC@NLO and WHIZARD of total FO cross sections at LO and NLO QCD for heavy vector boson (pair) production in association with light jets. Results are shown for a 13 TeV pp collider, with cuts applied as explained in section 4.2.1. Statisical errors of the integration, as well as the individual K-factor for each program, are shown. For better comparison, we provide the significance $\sigma^{\mathrm{sig}}_{\mathrm{NLO}}$ of the deviation between the NLO results in the last column. LO MG5_aMC@NLO results are mostly taken from [33], with an asterisk * marking recalculated values. NLO results are always recalculated, with the exception of $pp \to W^{\pm}jjj$, marked by a dagger †, which could not be recomputed with MG5_aMC@NLO for the reasons stated in section 4.2.2. Processes calculated in the four-flavor scheme are marked with $(4f)$.

Process	MG5_aMC@NLO			WHIZARD			
3 to 4 vector bosons plus jet	σ_{LO}[fb]	σ_{NLO}[fb]	K	σ_{LO}[fb]	σ_{NLO}[fb]	K	$\sigma^{\mathrm{sig}}_{\mathrm{NLO}}$
$pp \to W^+W^-W^\pm(4f)$	$1.307(3) \cdot 10^2$	$2.111(4) \cdot 10^2$	1.62	$1.3033(21) \cdot 10^2$	$2.1170(22) \cdot 10^2$	1.62	1.31
$pp \to W^+W^-W^\pm j(4f)$	$9.167(10) \cdot 10^1$	$12.00(4) \cdot 10^1$	1.31	$9.149(26) \cdot 10^1$	$11.985(29) \cdot 10^1$	1.31	0.30
$pp \to ZW^+W^-(4f)$	$9.658(65) \cdot 10^1$	$16.751(27) \cdot 10^1$	1.73	$9.742(15) \cdot 10^1$	$16.824(16) \cdot 10^1$	1.73	2.33
$pp \to ZW^+W^-j(4f)$	$8.340(10) \cdot 10^1$	$10.590(17) \cdot 10^1$	1.27	$8.325(23) \cdot 10^1$	$10.600(26) \cdot 10^1$	1.27	0.32
$pp \to ZZW^\pm$ *	$3.073(6) \cdot 10^1$	$5.630(12) \cdot 10^1$	1.83	$3.062(6) \cdot 10^1$	$5.615(6) \cdot 10^1$	1.83	1.12
$pp \to ZZW^\pm j$	$2.810(4) \cdot 10^1$	$3.652(7) \cdot 10^1$	1.30	$2.816(8) \cdot 10^1$	$3.684(9) \cdot 10^1$	1.31	2.81
$pp \to ZZZ$ *	$1.0833(18) \cdot 10^1$	$1.431(5) \cdot 10^1$	1.32	$1.0842(16) \cdot 10^1$	$1.4295(16) \cdot 10^1$	1.32	0.29
$pp \to ZZZj$	4.823(11)	6.386(23)	1.32	4.845(16)	6.391(18)	1.32	0.17
$pp \to W^+W^-W^\pm Z(4f)$ *	0.6575(12)	1.233(3)	1.88	0.6604(18)	1.2404(20)	1.88	2.05
$pp \to W^+W^-ZZ(4f)$	$4.320(13) \cdot 10^{-1}$	$7.100(13) \cdot 10^{-1}$	1.64	$4.331(12) \cdot 10^{-1}$	$7.163(12) \cdot 10^{-1}$	1.65	3.56
$pp \to W^\pm ZZZ$	$0.5862(10) \cdot 10^{-1}$	$1.2446(26) \cdot 10^{-1}$	2.12	$0.5869(8) \cdot 10^{-1}$	$1.2468(13) \cdot 10^{-1}$	2.12	0.76

Table 4.5. Comparison between MG5_aMC@NLO and WHIZARD of total FO cross sections at LO and NLO QCD for the production of three to four heavy vector bosons in association with light jets. Results are shown for a 13 TeV pp collider, with cuts applied as explained in section 4.2.1. Statisical errors of the integration, as well as the individual K-factor for each program, are shown. For better comparison, we provide the significance $\sigma^{\mathrm{sig}}_{\mathrm{NLO}}$ of the deviation between the NLO results in the last column. LO MG5_aMC@NLO results are mostly taken from [33], with an asterisk * marking recalculated values, while NLO results are always recalculated. Processes calculated in the four-flavor scheme are marked with $(4f)$.

Process	MG5_aMC@NLO			WHIZARD			
Jets, top quarks, and bosons	σ_{LO}[fb]	σ_{NLO}[fb]	K	σ_{LO}[fb]	σ_{NLO}[fb]	K	$\sigma^{\mathrm{sig}}_{\mathrm{NLO}}$
$pp \to jj$ *	$1.1593(23) \cdot 10^9$	$1.6040(29) \cdot 10^9$	1.38	$1.162(4) \cdot 10^9$	$1.601(5) \cdot 10^9$	1.38	0.10
$pp \to jjj$	$8.940(21) \cdot 10^7$	$7.619(19) \cdot 10^7$	0.85	$9.01(4) \cdot 10^7$	$7.46(9) \cdot 10^7$	0.83	1.73
$pp \to t\bar{t}$	$4.584(3) \cdot 10^5$	$6.746(14) \cdot 10^5$	1.47	$4.589(9) \cdot 10^5$	$6.740(10) \cdot 10^5$	1.47	0.35
$pp \to t\bar{t}j$ *	$3.133(5) \cdot 10^5$	$4.095(8) \cdot 10^5$	1.31	$3.123(6) \cdot 10^5$	$4.087(9) \cdot 10^5$	1.31	0.66
$pp \to t\bar{t}jj$ *	$1.363(3) \cdot 10^5$	$1.784(3) \cdot 10^5$	1.31	$1.360(4) \cdot 10^5$	$1.775(7) \cdot 10^5$	1.31	1.18
$pp \to t\bar{t}t\bar{t}$	4.505(5)	9.076(13)	2.01	4.485(6)	9.070(9)	2.02	0.38
$pp \to t\bar{t}W^{\pm}$	$3.777(3) \cdot 10^2$	$5.668(18) \cdot 10^2$	1.50	$3.775(5) \cdot 10^2$	$5.674(5) \cdot 10^2$	1.50	0.32
$pp \to t\bar{t}W^{\pm}j$ *	$2.352(3) \cdot 10^2$	$3.434(8) \cdot 10^2$	1.46	$2.356(7) \cdot 10^2$	$3.427(8) \cdot 10^2$	1.45	0.62
$pp \to t\bar{t}Z$	$5.273(4) \cdot 10^2$	$7.652(25) \cdot 10^2$	1.45	$5.260(7) \cdot 10^2$	$7.639(9) \cdot 10^2$	1.45	0.49
$pp \to t\bar{t}Zj$	$3.953(4) \cdot 10^2$	$5.079(14) \cdot 10^2$	1.28	$3.943(14) \cdot 10^2$	$5.069(17) \cdot 10^2$	1.29	0.45
$pp \to t\bar{t}W^+W^-(4f)$	6.675(6)	9.934(19)	1.49	6.679(9)	9.943(11)	1.49	0.41
$pp \to t\bar{t}ZZ$	1.349(14)	1.843(4)	1.37	1.3590(29)	1.842(3)	1.36	0.20
$pp \to HZ$	$6.468(8) \cdot 10^2$	$7.693(19) \cdot 10^2$	1.19	$6.474(11) \cdot 10^2$	$7.679(12) \cdot 10^2$	1.19	0.62

Table 4.6. Comparison between MG5_aMC@NLO and WHIZARD of total FO cross sections at LO and NLO QCD for the production of top quarks and/or light jets in association with heavy gauge bosons. Results are shown for a 13 TeV pp collider, with cuts applied as explained in section 4.2.1. Statisical errors of the integration, as well as the individual K-factor for each program, are shown. For better comparison, we provide the significance $\sigma^{\mathrm{sig}}_{\mathrm{NLO}}$ of the deviation between the NLO results in the last column. LO MG5_aMC@NLO results are mostly taken from [33], with an asterisk * marking recalculated values, while NLO results are always recalculated. Processes calculated in the four-flavor scheme are marked with $(4f)$.

4.3. Comparison of Fixed-order Differential Distributions

Having validated our results for the total cross sections, we turn our attention to the validation of fNLO QCD event generation with WHIZARD, as described in section 3.1.5. As stated therein, FO events are sets of four-momenta, associated with a weight, and cannot fully describe experimental data without further showering and hadronization. Nevertheless, FO differential distributions pose a physically meaningful description of perturbative QCD up to the chosen order in α_S. The differential distributions of an IR-safe observable computed from the sets of FO events generated by different MC event generators should therefore agree within the (combined) statistical and theoretical uncertainties.

In this section we will compare the results for FO differential distributions of events generated by WHIZARD with those generated by MG5_aMC@NLO (in version `2.7.3`), which was introduced in section 4.2. In addition, we also compare against the results of the multi-purpose MC event generator Sherpa [34, 175] (in version `2.2.10`), which uses the CS subtraction scheme [40] for its NLO implementation. This provides additional cross-scheme validation.

Our focus in this section and beyond will be on processes with n-jet final states at a future lepton collider such as the ILC. There, these processes will play an important role not only as QCD corrections to the process $e^+e^- \to f\bar{f}$, but also as a relevant background to e.g. multi-boson processes with hadronic final states. Moreover, these multi-jet processes provide an excellent opportunity to conduct high-precision studies on e.g. hadronization models or SM parameters like the value of α_S. Deviations from these parameters can provide signals in searches for BSM physics.

While there have been many studies on multi-jet processes in electron-positron annihilation at NLO QCD (e.g. [81–84]), and up to NNLO for up to three light jets (e.g. [185, 186]), even more recent studies at NLO (e.g. [187]) often focus on LEP CM energies or the lower end of proposed energy ranges at future lepton colliders. With discussed CM energies of the first stage of the ILC of 250 GeV to 500 GeV [188], the technical design report [21] also includes future upgrade paths up to 1 TeV. Thus, a comparison between MC event generators at this energy for exclusive n-jet observables remains relevant for scientific discussion.

Additionally, since the FKS implementation in WHIZARD also supports hadronic initial states, we chose the process $pp \to Zj$ with an on-shell Z boson in order to validate and highlight the fNLO QCD event generation capabilities of WHIZARD for hadron colliders.

We will first give an overview over the process and parameter setup for the event

generation in section 4.3.1, followed by the details for the setup of the analyses in section 4.3.2, where we also explain the definitions of the event-shape observables we considered. The results for n-jet processes in lepton collisions are then split into differential distributions of observables defined on clustered jets presented in section 4.3.3 and differential distributions of event-shape observables presented in section 4.3.4. The results for the hadron-collider process $pp \to Zj$ are briefly discussed in section 4.3.5.

4.3.1. Process and Parameter Setup for Event Generation

For the event generation with all three programs we mostly used the same settings and parameters as already described in section 4.2.1. There are, however, a few differences and additional settings, which will be explained in the following.

Beginning with the central scale $\mu = \mu_R = \mu_F$, we chose a fixed scale

$$\mu = m_Z \tag{4.3.1}$$

simply because of the fact that the implementation of the running of the strong coupling is different between Sherpa, MG5_aMC@NLO, and WHIZARD. The former program uses the Landau pole Λ as a reference, thus leading to logarithms of $\log(\mu_R^2/\Lambda^2)$, while the latter two rely on the physical mass m_Z of the Z-boson, which leads to logarithms of $\log(\mu_R^2/m_Z^2)$. For processes with high orders of α_S, as well as a running of α_S beyond one-loop order in the β-function, the influence of the different approaches can be seen as a significant deviation between total cross sections already at Born-level, when only taking into account statistical uncertainties. Certainly, when also considering theoretical uncertainties in the form of scale variations, the significance of this difference will vanish. Since we are mostly interested in a direct comparison of the capabilities of our implementation of FKS, however, we decided to minimize uncertainties stemming from different implementations of the running of α_S by choosing a fixed scale.[2] Nevertheless, we additionally generated event samples with WHIZARD for typical scale variations of 2μ and $\mu/2$ in order to put the statistical errors into perspective.

The next change in our setup is that instead of the anti-k_t algorithm we chose the generalized k_t jet-clustering algorithm for e^+e^- collisions [162] for all lepton-collider processes considered in this section. It is inspired by the original k_t algorithms [189], with extensions akin to those in [181] for the generalized pp algorithm. In contrast to many other k_t jet algorithms, the generalized e^+e^- k_t algorithm incorporates spherical

[2] WHIZARD actually does offer the possibility to use either the Landau Pole or the mass of the Z-boson as a reference in the running of α_S. However, MG5_aMC@NLO and Sherpa differ in their implementation and do not offer both options.

instead of cylindrical coordinates. The reason for the use of spherical coordinates is that at a lepton collider the full kinematics of an event can be reconstructed, whereas at a hadron collider cylindrical coordinates are preferred due to the fact that only transverse kinematics can be fully reconstructed there.

In the following, we adopt the definitions and explanations from [162] for the generalized e^+e^- k_t algorithm, where – for the reasons outlined above – the distance measures of the generalized e^+e^- k_t algorithm are defined as

$$d_{kl} = \min\left(E_k^{2p}, E_l^{2p}\right) \frac{1-\cos\theta_{kl}}{1-\cos R}\,, \tag{4.3.2}$$

$$d_{kB} = E_{kB}^{2p}\,, \tag{4.3.3}$$

with R being the real-valued jet-radius parameter and p an arbitrary real number. As for all k_t jet algorithms, the minimum $d_{\min} = \min\left(d_{kl}, d_{kB}\right)$ of all distance measures d_{kl} between each pair of final-state particles k and l, and all distance measures d_{kB} between each particle k and the beam axis, is recursively determined. For each new $d_{\min}$ that is found to be a d_{kl}, particles k and l are merged into a new particle (also called pseudojet), which receives the sum of their four-momenta as its momentum.[3] If $d_{\min} = d_{kB}$ we call k an *inclusive* jet. The algorithm continues until only inclusive jets remain in the event.

With the restriction $R \leq \pi$, the behavior of the algorithm is analogous to the hadron-collider algorithms in the way that an inclusive jet is produced by a particle or pseudojet with index k if its angle θ_{kX} to all other objects X meets the criteria $\theta_{kX} > R$.[4] We chose

$$R = 0.5\,, \quad p = -1\,, \tag{4.3.4}$$

which is analogous to the anti-k_t jet algorithm and to the setting for R we used in section 4.2. The above choice of R and p for the generalized e^+e^- k_t algorithm leads to an IR-safe lepton-collider jet algorithm, with energy-hard jets that form well-separated cone-like objects, having $2R$ as the opening angle. Thus, its behavior is effectively similar to cone algorithms, where particles within a specific angular region of conical shape are defined as a jet.

At the time of writing this thesis, only WHIZARD supports the generalized e^+e^- k_t algorithm natively. For both MG5_aMC@NLO and Sherpa, the jet definition had to be changed in the code of their respective FastJet interfaces.

Another notable difference in the setup included the change from using the separate

[3] This is the so-called E-scheme recombination, with other schemes existing as well.

[4] For $R > \pi$, the denominator of eq. (4.3.2) is replaced by $3 + \cos R$. This requires the notion of *exclusive* jets, as for $\pi < R < 3\pi$ all final-state particles will be merged into a single inclusive jet. For more information see [162].

integration of components in WHIZARD, as used for the validation of the total cross sections in section 4.2, to using the combined integration. Firstly, this serves as another validation of a different feature, and secondly it vastly simplifies the event generation with WHIZARD, as the events for the separate components do not have to be simulated via individual runs of the program (see section 3.1.5) and do not have to be merged afterwards in the analysis.

We translated the whole process setup into the Sherpa run card syntax. Like WHIZARD, Sherpa allows the user to fix the order of the EW coupling to the lowest order possible (see fig. 2.1 and section 4.2.1). An example for a Sherpa run card is given in appendix F.2 for the process $e^+e^- \to jjj$. For the finite virtual one-loop contributions, we used the OpenLoops interface provided by Sherpa.

With the described setup, we generated weighted FO events at NLO QCD for the processes listed in table 4.7 with all three MC event generators. There, we also state the number of generated fNLO event groups (see section 3.1.5). Due to the differences both in the applied subtraction scheme and in the technical implementations, the actual number of total (sub)events can vary significantly across the programs. We also note that, in contrast to Sherpa and WHIZARD, in MG5_aMC@NLO the number of FO event groups cannot be requested directly and is indirectly influenced by the requested integration accuracy. For this reason, the number of event groups simulated by MG5_aMC@NLO can in general differ from Sherpa and WHIZARD. For jet multiplicities of four and five jets, however, they turned out to be identical.

In order to put the results into perspective and to give a better understanding of the effects of NLO corrections, we generated weighted FO events at LO with WHIZARD as well, where the number of generated events is chosen to be the same as the generated event groups at NLO.

Process	MG5_aMC@NLO	Sherpa	WHIZARD
$e^+e^- \to jj$	$\sim 0.2 \cdot 10^8$	$1 \cdot 10^8$	$1 \cdot 10^8$
$e^+e^- \to jjj$	$\sim 1.22 \cdot 10^8$	$1 \cdot 10^8$	$1 \cdot 10^8$
$e^+e^- \to jjjj$	$5 \cdot 10^8$	$5 \cdot 10^8$	$5 \cdot 10^8$
$e^+e^- \to jjjjj$	$5 \cdot 10^8$	$5 \cdot 10^8$	$5 \cdot 10^8$
$pp \to Zj$	$\sim 0.93 \cdot 10^8$	$1 \cdot 10^8$	$1 \cdot 10^8$

Table 4.7. The Number of FO event groups generated at NLO QCD per process and MC event generator are shown. See section 4.3.1 for details. For WHIZARD, as many LO events have been generated as for NLO.

Cross-check via Total Cross Sections

We already validated the capability of WHIZARD to compute fNLO QCD cross sections reliably in section 4.2 against MG5_aMC@NLO. Therefore, we can now cross-check our event generation setup among the programs before actually simulating events by comparing the computed total cross sections in table 4.8. We see a perfect agreement within the statistical errors.

	WHIZARD	MG5_aMC@NLO		Sherpa	
Process	σ_{NLO}[fb]	σ_{NLO}[fb]	$\sigma^{\text{sig}}_{\text{NLO}}$	σ_{NLO}[fb]	$\sigma^{\text{sig}}_{\text{NLO}}$
$e^+e^- \to jj$	646.67(7)	646.69(4)	0.25	646.621(27)	0.18
$e^+e^- \to jjj$	274.1(7)	274.3(5)	0.23	274.2(5)	0.12
$e^+e^- \to jjjj$	75.36(28)	75.32(15)	0.13	75.33(10)	0.10
$e^+e^- \to jjjjj$	15.31(7)	15.33(3)	0.26	15.27(5)	0.47
$pp \to Zj$	$1.0180(21) \cdot 10^7$	$1.0186(9) \cdot 10^7$	0.26	$1.0171(10) \cdot 10^7$	0.39

Table 4.8. The total fNLO QCD cross sections for the processes considered in section 4.3 for each MC event generator used in the comparison of differential distributions are shown. Results have been calculated for fixed scales $\mu = m_Z$ (see section 4.3.1 for detailed setup) and are shown with statistical errors. Additionally, the significance $\sigma^{\text{sig}}_{\text{NLO}}$ (see eq. (4.2.8)) of the deviation between the WHIZARD result and that of MG5_aMC@NLO and Sherpa, respectively, is shown.

4.3.2. Setup of the Analyses

All three programs, MG5_aMC@NLO, Sherpa, and WHIZARD, come with their own analysis tools. However, using the same analysis framework with exactly the same analysis for each MC event generator eliminates sources of errors and streamlines the workflow. For this reason, in order to analyze the generated FO event samples, we used the Rivet toolkit [190]. It provides a well established and validated framework to easily write highly flexible and customizable analyses in the form of cleanly-structured C++ plugins. For this purpose, it comes with a vast library of readily available functions to compute differential distributions of common observables used in both lepton and hadron-collider communities. User-defined features and observables can be simply implemented thanks to the modular nature of the plugin approach. Rivet also provides an interface to FastJet [162] to apply different jet-clustering algorithms. Event samples are read in the HepMC [141, 142] format, with the analysis output being written as Yoda [191] files to be plotted via scripts provided by Rivet.

While WHIZARD provides the option to write event samples in the HepMC format and Sherpa even comes with a direct Rivet interface, MG5_aMC@NLO

unfortunately does not have the option to write FO events as HepMC files. It can use the LHEF format, however, which is easily converted to the HepMC format via the `convert_example.exe` script provided by HepMC. From HepMC 3.2.3 onward, this also correctly takes into account correlated event groups present in fNLO event records.

As mentioned in section 3.1.5, these fNLO event groups have to be handled with care when filling their subevents' weights into binned data types such as histograms, in order to ensure the correct statistical treatment. For our analyses, we used Rivet 2, which does not have a native treatment for fNLO event groups. However, thanks to the modular plugin approach, it is very straightforward to implement a rudimentary solution in the way described at the end of section 3.1.5.

Indeed, in the Rivet `Contrib` suite [192] there already exists a plugin in the form of the `NLOHisto1D` class. By using this class, for each event group a temporary histogram is filled such that real and counterevent weights in the collinear and/or soft limit in almost all cases are filled into the same bin with their weights added. The temporary bins filled this way, each with their own sum of weights, are then added to the respective bins of the final histogram, with each non-trivial temporary bin counting effectively as a single event and the square of the sum of weights of the temporary bin is added to the sum of squared weights of the final bin. This approach does not alleviate misbinning effects as described in the end of section 3.1.5.[5] However, the quality of the produced histograms is not reduced by these effects, especially since misbinning does not occur very often.

It is important to note that while Rivet 3 has a native support for correlated event groups, at the time of writing this thesis there is an ongoing discussion about the details regarding the treatment of statistical errors by Rivet 3 in the analysis of fNLO event groups generated by WHIZARD. It might be possible that the method described above with the `NLOHisto1D` class and Rivet 2 actually underestimates the statistical errors of single bins to some degree, since the total weight of the event group is split among the subevents of that group.

However, the subevents can end up in different bins, reducing the effective sample size, which should in turn *increase* the statistical error in their respective bins. In Rivet 3, in order to account for this effect, each subevent is counted as a fraction of the total number of subevents in an event group and this fraction is then used as a weighting factor in the calculation of the sum of squared weights (see [173, appendix A]). This, however, might be in conflict with how event weights in WHIZARD are normalized at fNLO, which is one of the reasons for why we decided on using the more "naive" treatment by the `NLOHisto1D` class together with Rivet 2 until a conclusion is reached

[5] Rivet 3 does implement a weight "smearing" to treat misbinning effects (see Appendix A in [173]).

in this regard.

For now, the statistical errors presented in this work have to be considered with this fact in mind. The overall conclusions drawn in the comparison between the different MC event generators is not affected by this, however, since the effects are negligible when also considering theoretical uncertainties in the form of scale variations, as can be seen in the results presented later in this section.

An example of an analysis used for the process $e^+e^- \to jjj$ is given in appendix F.3, with the analyses for the other considered processes being analogous. The general structure is always the same:

- use FastJet to construct inclusive jets, with the same jet definition as for the integration/simulation,
- use the same cuts and selection criteria as for the integration/simulation,
- use a jet-p_T ordering from p_T-hardest to -lowest jet for observables defined on jets (see text below for reasoning),
- scale the final histograms by a factor of the total cross section over the total sum of weights, σ_NLO/SoW.

Regarding the ordering of the jets by their transverse momentum mentioned above, we note that for lepton-collider events that are truly at rest the natural choice would be to order them by their energies, as this carries more information. However, events can only be truly at rest at the absence of any beamstrahlung and without ISR of photons. While we have neglected such effects in our comparison study and thus produced events without any longitudinal boost, in the real-world experiments these effects can only be neglected near LEP energies at the Z-pole, for example.[6] For a future lepton collider with CM energies well above the Z-pole, however, beamstrahlung and ISR effects will play a significant role. They can cause radiative returns to the Z-pole, in which case an ordering by jet-p_T can be beneficial, especially since this would preserve invariance under a longitudinal boost along the z-axis. Thus, at future experiments one will have to determine on a case-by-case basis – by evaluating theory and jet response uncertainties – which jet ordering will give the most information about the events that are studied [194].

In the following, we will give an overview and a short explanation of the observables considered in the comparison of differential distributions between MG5_aMC@NLO, Sherpa, and WHIZARD. Besides well-known observables of final state objects, such as the transverse momentum p_T, the energy E, the pseudorapidity η, and the invariant mass M_X of the system X made up by a subset of final-state objects, we also compared

[6] Similar conditions existed at the Stanford Linear Collider [193].

differential distributions of event-shape observables. As their name suggests, these observables are a means to quantify the topology of an event in terms of its jets in hadronic final states. However, when defining event-shape observables, we do not involve jet algorithms to associate particles with a jet, but rather consider all final-state particles of an event directly in order to determine a scalar value classifying the jet topology of the event.

Typically, event-shape observables are defined in the center-of-mass system and constructed in such a way that they vanish for a final state consisting of two particles (or two extremely narrow jets) that are perfectly back-to-back. While distributions of event-shape observables with peaks very close to zero will usually still indicate a dominance of dijet events, a peak at increasingly higher event-shape values can be interpreted as an increase in additional jets for most event-shape observables. Therefore, these observables are sensitive to the strong coupling constant α_S, and thus provide a possibility to determine the value of α_S via fits.

Further use cases are the measurement of the QCD color factors C_A and C_F (e.g. [195]), and the tuning of MC event generation aspects such as parton showers. Also, event shapes have played an important role both in developing and testing of analytic hadronization models (e.g. [196–198]). Last, but not least, event shapes can help in distinguishing between different processes and thus are important tools not only in precision studies, but can also help to separate signals from background noise in searches for BSM physics. There is a caveat for higher jet multiplicities, however, as for most event-shape observables, an increase in the number of jets in an event often shows an asymptotic behavior towards the maximum value of the observable in question. This can limit the distinguishing powers at high multiplicities for exclusive numbers of jets.

While there is a whole palette of different event-shape observables, we limited our comparison to a few, more common ones, explained in the following.

Thrust, Thrust Major, and Thrust Minor

The thrust event-shape observable T is defined as [75, 76]

$$T \equiv \max_{\mathbf{n}} \left(\frac{\sum_i |\mathbf{p}_i \cdot \mathbf{n}|}{\sum_i |\mathbf{p}_i|} \right) , \tag{4.3.5}$$

where we sum over all final state particles, each having a three-momentum of $\mathbf{p}_i$. The direction of the unit vector $\mathbf{n}$ that maximizes T then defines the thrust axis $\mathbf{n}_T$. With this definition, thrust describes how "pencil-like" or spherical an event is, as a dijet event that is perfectly back-to-back (resembling a pencil) would result in a thrust value of $T = 1$. In contrast, for an increasing number of particles, the more they

are spatially evenly distributed, the closer the thrust becomes to $T = 1/2$, which would resemble a perfectly-spherically symmetric distribution of particles. Because this behavior is opposite to the common construction of event-shape observables to vanish in the perfectly back-to-back dijet case, the thrust observable is usually applied as $1 - T$.

We can use eq. (4.3.5) to define the thrust major T_{M} with the thrust major axis $\mathbf{n}_{T_{\mathrm{M}}}$ that maximizes T_{M} in the plane perpendicular to $\mathbf{n}_T$ by imposing the constraint $\mathbf{n}_{T_{\mathrm{M}}} \cdot \mathbf{n}_T = 0$. Similarly, we can define the thrust minor axis $\mathbf{n}_{T_{\mathrm{m}}}$ by imposing it being perpendicular to both $\mathbf{n}_T$ and $\mathbf{n}_{T_{\mathrm{M}}}$, meaning $\mathbf{n}_T \times \mathbf{n}_{T_{\mathrm{M}}} = \mathbf{n}_{T_{\mathrm{m}}}$. The associated thrust minor T_{m} is then simply obtained by the expression in parentheses in eq. (4.3.5) with $\mathbf{n} = \mathbf{n}_{T_{\mathrm{m}}}$. Resulting from the three different thrust definitions is an ordering of $T \geq T_{\mathrm{M}} \geq T_{\mathrm{m}}$.

Both thrust major and minor vanish for a back-to-back dijet event. However, T_{M} already becomes non-zero for deviations from a pencil-like event shape, whereas T_{m} becomes non-zero only for 4-jet final states and beyond with typical values of $T_{\mathrm{m}} \leq 1/2$. By *typical values* we mean the values where the bulk of the distribution lies and we quote them from [199]. For the thrust major, these values are $T_{\mathrm{M}} \leq 1/3$ for 3-jet final states and $T_{\mathrm{M}} \leq 1/\sqrt{2}$ the more spherical an event becomes.

Using both the definitions of thrust major and minor, we can then define the oblateness O of an event by simply taking [200]

$$O \equiv T_{\mathrm{M}} - T_{\mathrm{m}} \,. \tag{4.3.6}$$

It vanishes both for back-to-back dijet and spherical n-jet events because in these cases the momentum distribution becomes isotropic in the plane defined by the thrust major and minor axes, leading to a decreasing difference between T_{M} and T_{m}. Typical values for oblateness between these two extrema are $O \leq 1/3$.

Hemisphere Masses and Broadenings

The definition of the thrust axis $\mathbf{n}_T$ allows us to define a plane perpendicular to it, dividing the event into two hemispheres S_+ and S_-, with the invariant masses M_{S_+} and M_{S_-} of all particles in each hemisphere, respectively. The high hemisphere mass M_{high} is then defined as [201, 202]

$$M_{\mathrm{high}} \equiv \max\left(M_{S_+}, M_{S_-}\right), \tag{4.3.7}$$

and the low hemisphere mass as the minimum. It is common to scale the hemisphere masses by the center-of-mass energy $\sqrt{s}$ or the visible energy E_{vis} of an event. In our comparison we chose the definition of the scaled high hemisphere squared mass $M^2_{\mathrm{high,s}}$

available in RIVET as

$$M^2_{\text{high,s}} \equiv \frac{M^2_{\text{high}}}{E^2_{\text{vis}}} = \frac{1}{E^2_{\text{vis}}} \max\left(\left| \sum_{\mathbf{p}_i \cdot \mathbf{n}_T > 0} p_i \right|^2 , \left| \sum_{\mathbf{p}_i \cdot \mathbf{n}_T < 0} p_i \right|^2 \right) , \tag{4.3.8}$$

with final-state momenta p_i. For massless partons, a perfectly back-to-back dijet event would again have $M^2_{\text{high,s}} = 0$, and a perfectly spherical event would equate to a value of $M^2_{\text{high,s}} = 1/2$ for $E^2_{\text{vis}} = s$, whereas e.g. a trijet event would lead to typical values of $M^2_{\text{high,s}} \leq 1/3$.

While the event-shape observables presented so far have been mostly longitudinal in nature, we can also define a measure for the transverse extension of an event in each hemisphere. These so-called hemisphere (or jet) broadenings $B_\pm$ are defined as [203]

$$B_{S_\pm} \equiv \left(\sum_{\pm \mathbf{p}_i \cdot \mathbf{n}_T > 0} |\mathbf{p}_i \times \mathbf{n}_T| \right) \left(2 \sum_i |\mathbf{p}_i| \right)^{-1} . \tag{4.3.9}$$

The total hemisphere broadening B_{tot} and wide hemisphere broadening B_{w} are subsequently defined as

$$B_{\text{tot}} \equiv B_{S_+} + B_{S_-} , \quad B_{\text{w}} \equiv \max\left(B_{S_+}, B_{S_-} \right) . \tag{4.3.10}$$

Both observables again evaluate to zero for a perfectly back-to-back dijet event. Deviating from this event shape and adding another jet increases both equally to typical values of up to $1/(2\sqrt{3})$. While higher jet multiplicities do not change this upper limit for B_{w}, typical values for the total jet broadening can increase to $B_{\text{tot}} \leq 1/(2\sqrt{2})$ for events with high jet multiplicities and spherically-evenly distributed momenta.

Parisi Event-shape Variables

The Parisi event-shape tensor $\theta^{\alpha\beta}$ is a linearized version of the sphericity tensor [204] and thus is an IR-safe observable. It is defined as [77, 78]

$$\theta^{\alpha\beta} \equiv \left(\sum_i \frac{p_i^\alpha p_i^\beta}{|\mathbf{p}_i|} \right) \left(\sum_i |\mathbf{p}_i| \right)^{-1} , \quad \alpha, \beta \in \{1, 2, 3\} , \tag{4.3.11}$$

with a normalization that is chosen such that the three eigenvalues λ_i sum up to unity, $\lambda_1 + \lambda_2 + \lambda_3 = 1$. Reducing the tensor θ to a diagonal one then gives the characteristic equation [79]

$$\lambda^3 - \lambda^2 + \frac{1}{3} C \lambda - \frac{1}{27} D = 0 , \tag{4.3.12}$$

from which λ_i can be obtained. The Parisi event-shape variables C and D consequently turn out to be

$$C = 3\left(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3\right) , \quad D = 27\lambda_1\lambda_2\lambda_3 . \tag{4.3.13}$$

These variables both vanish for a back-to-back dijet event. In case the event is planar in shape, D still vanishes, while [79]

$$C = 3\lambda_1\left(1 - \lambda_1\right) . \tag{4.3.14}$$

The variable D only becomes non-zero for events deviating from a planar shape. Therefore, C can be seen as a measure of the planarity of an event, while D is a measure for the spherical distribution of the event.

Resummation of Large Logarithms

Observables that are not inclusive over the full phase space – e.g. due to cuts or jet selection criteria – will exhibit logarithmic enhancements at every order in the perturbation series. The reason for this is that in this case the cancellation of IR divergences between the real and virtual contribution is not exact anymore. This leads to logarithms of large fractions of different (momentum) scales. For example, for an IR-safe event-shape observable y that vanishes for a pencil-like event (i.e. in the dijet limit), the perturbation series of its distribution at $\mathrm{N}^k\mathrm{LO}$ turns out to be [205]

$$\frac{1}{\sigma}\frac{d\sigma_k}{dy} \sim \alpha_S^k \frac{1}{y}\log^{2k-1}\frac{1}{y} , \quad \text{for } y \ll 1 . \tag{4.3.15}$$

As a consequence, in order for the perturbation series to converge, it is not sufficient to have $\alpha_S \ll 1$ anymore and we require $\alpha_S \log^2 y \ll 1$ [112]. Unfortunately, the bulk of the distribution usually lies at small values of y. In order to increase the predictive power of perturbative computations, these large logarithms have to be resummed at all orders of the perturbation series. The obtained resummed expression then has to be matched to the FO results.

In this work, we focused on the pure fNLO corrections in QCD. Therefore, in the results we present throughout this thesis, logarithmic enhancements will affect the K-factor (see eq. (4.2.7)) in addition to kinematic effects due to the real emission. In the discussion of our results, however, we will focus on the kinematic effects.

A Note on Scale Variations and Negative Bin Weights

As mentioned in the discussion on fNLO events in WHIZARD (see section 3.1.5), differential distributions of fNLO observables can have bins with a negative sum of

weights, which can be circumvented by choosing bin widths such that subevents with positive weights outweigh the negative ones. Finding the correct bin width without concealing interesting parts of the distribution is very tedious, however, and thus we limited our efforts in this regard to the results for the central scale μ throughout the thesis. As a consequence, scale variations of the central result with WHIZARD might occasionally show large excursions to negative values in regions where e.g. Born-like subevents predominate, simply because the negative weight values for these subevents become larger with a different scale. This effect is rather rare, though, and when it happens the statistical error alone will be sufficient to assess the agreement between MG5_aMC@NLO, Sherpa, and WHIZARD.

4.3.3. Differential Distributions of Jet Observables in $e^+e^- \to n$ Jets

In order to validate NLO effects of FSR, we first examine NLO differential distributions of pure n-jet final states in e^+e^- collisions. In this section we will focus on observables defined on jets clustered by the generalized e^+e^- k_t algorithm, namely jet momenta p_{T,j_i}, jet energy E_{j_i}, absolute[7] jet pseudorapidity $|\eta_{j_i}|$, and the invariant mass $M_{j_1 j_2}$ of the system of the two p_T-hardest jets. The results for the various event-shape observables with a direct comparison between jet multiplicities for each observable are then given in section 4.3.4 to better visualize their behavior.

To avoid repeating the same observation for the majority of histograms that are shown and discussed in the following, we can preemptively conclude that differential distributions at fNLO QCD obtained by analyzing events generated with WHIZARD show a very good agreement with results from MG5_aMC@NLO and Sherpa. This is not only true within the generally larger theoretical uncertainties estimated by scale variations, but mostly also with respect to the much smaller statistical MC errors. Furthermore, deviations between the central values are at the order of a few percent at most. Albeit sparsely needed, a more detailed discussion for differences between MG5_aMC@NLO, Sherpa, and WHIZARD will be given whenever necessary.

Differential Distributions for $e^+e^- \to jj$

In order to compare effects of FSR of a gluon from final-state quarks in lepton collisions, we start with the simple dijet process. The transverse momentum $p_{\mathrm{T},j_{1/2}}$ of the p_T-hardest and second-hardest jet j_1 and j_2, respectively, are shown in fig. 4.1. In both cases we see a rising curve for higher jet-p_T, the kind of continuum we would expect

[7] We checked that the pseudorapidity distribution is symmetric, which allows us to increase statistics by using the absolute value.

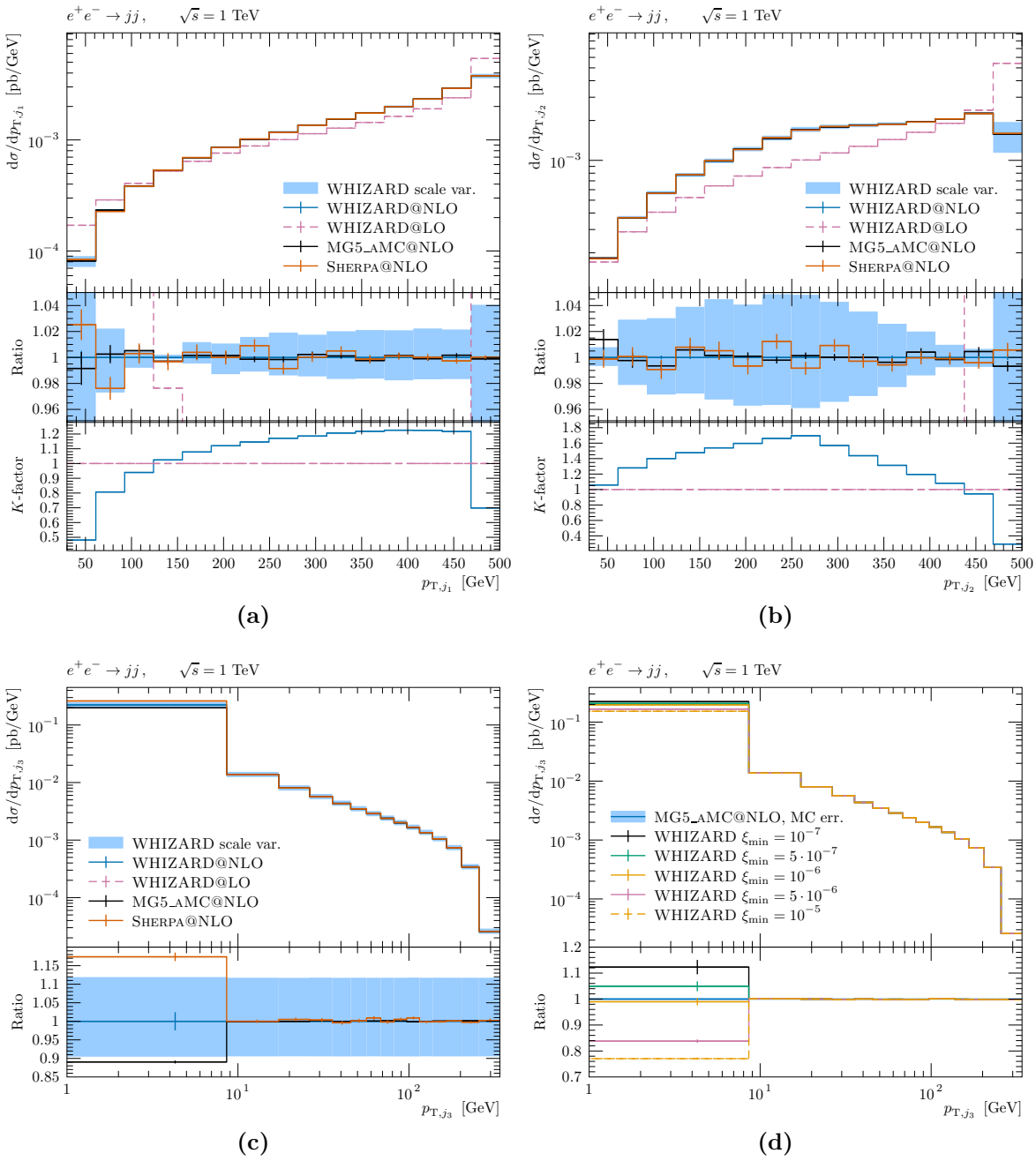


Figure 4.1. Differential distributions of jet transverse momenta p_{T,j_i} for the process $e^+e^- \to jj$ at fNLO QCD of the first- (a), second- (b), and third-hardest (c) jet. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result. In (d), the impact of varying $\xi_{\min}$ in WHIZARD on the p_{T,j_3} distribution is compared to the MG5_aMC@NLO result. There, the (tiny) error band for MG5_aMC@NLO represents statistical MC errors, not scale variations. In this case the ratio is taken with respect to the MG5_aMC@NLO results.

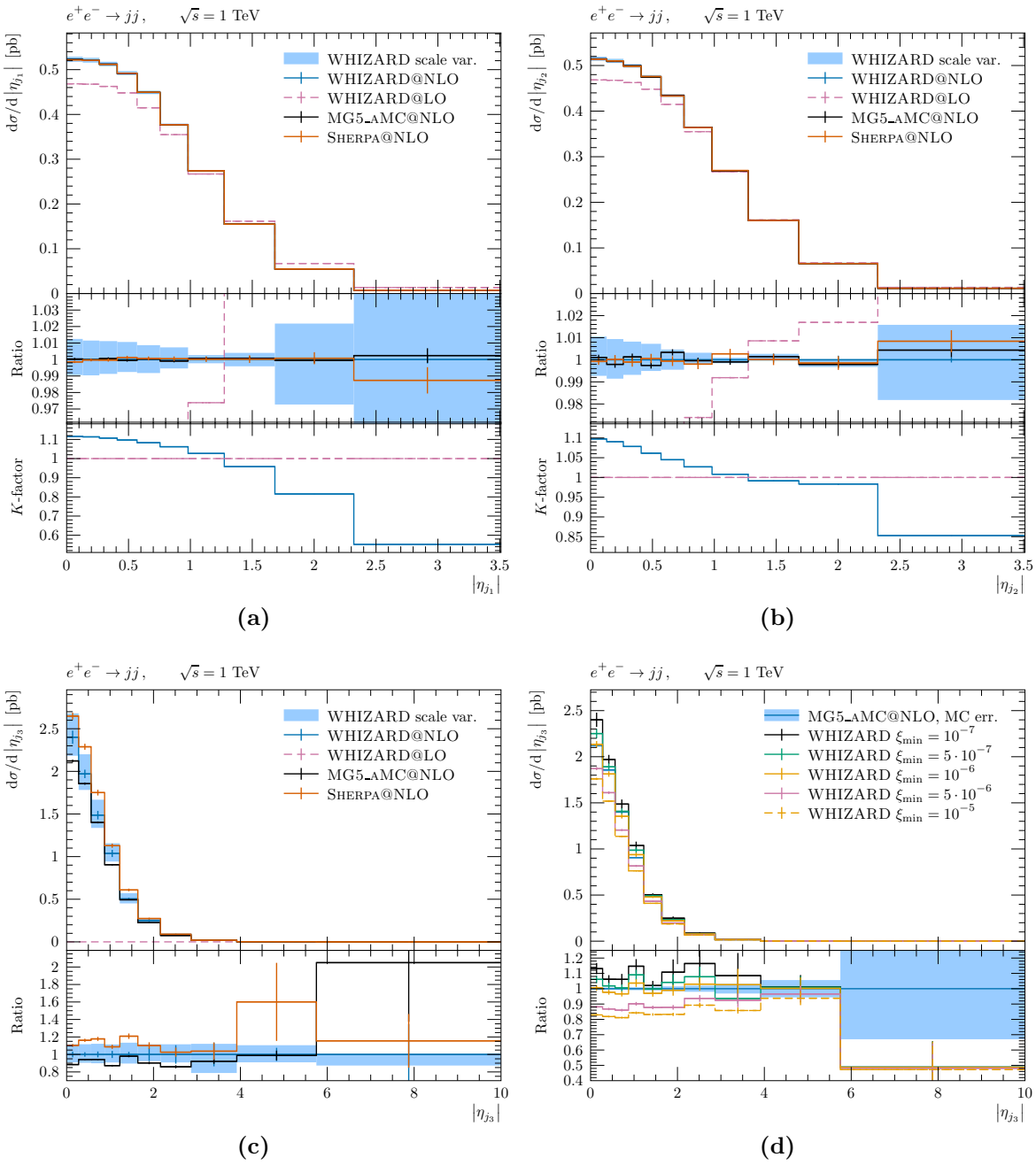


Figure 4.2. Differential distributions of jet pseudorapidities $|\eta_{j_i}|$ for the process $e^+e^- \to jj$ at fNLO QCD of the first- (a), second- (b), and third-hardest (c) jet. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result. In (d), the impact of varying $\xi_{\min}$ in WHIZARD on the $|\eta_{j_3}|$ distribution is compared to the MG5_aMC@NLO result. There, the error band for MG5_aMC@NLO represents statistical MC errors, not scale variations. In this case the ratio is taken with respect to the MG5_aMC@NLO results.

for FO dijet events at a lepton collider. Due to the NLO real corrections of an emitted gluon, compared to the LO results the peak of the distribution is significantly suppressed, with more jets shifted to lower p_{T}. By the nature of our chosen p_{T}-ordering, this effect is more prominent for the second-hardest jet. At the low end of the p_{T} spectrum, however, the NLO p_{T,j_1} distribution is significantly lower than at LO because soft radiation makes up the majority of low-p_{T} jets in this phase-space region.

We include the p_{T,j_3} distribution of the third-hardest jet j_3 in fig. 4.1 as well. It is strictly speaking only of LO accuracy due to the fact that this plot only represents the case of real radiation with a 3-jet final state, while counterevents are constructed for unresolved real partons and thus have a 2-jet final state. Hence, the p_{T,j_3} distribution effectively is that of the third-hardest jet of the LO process $e^+e^- \to jjj$ if no cuts were imposed on j_3. Consequently, all events contributing to this distribution are of $\mathcal{O}(\alpha^2\alpha_S)$, which is directly reflected in the trivial scale-variation band.

Because of the fact that we did not impose any cuts on the third jet, the p_{T,j_3} distribution can also be regarded as "unphysical" due to the unregulated IR divergence for small p_{T}, which can be seen in the first bin of fig. 4.1(c).[8] Nevertheless, showing this plot here allows us to highlight how differences in the subtraction scheme implementation of the different event generators can manifest. As shown in fig. 4.1(c), there is a slight deviation between the three event generators in the lowest-p_{T}, IR-divergent bin. This is not due to statistical fluctuations, as one might suspect, but rather due to small arbitrary cuts on the FKS variable ξ (or any equivalent implementation for CS subtraction in Sherpa) in order to prevent numerical instabilities in the soft limit $\xi \to 0$. In WHIZARD, for example, we employ a default cut of $\xi_{\mathrm{min}} = 10^{-7}$. Comparing different values of ξ_{min} to the MG5_aMC@NLO results in fig. 4.1(d), we find a good agreement for $\xi_{\mathrm{min}} = 10^{-6}$.

We refrain from changing the default of ξ_{min} in WHIZARD, however, since the importance sampling of ξ differs significantly between MG5_aMC@NLO and WHIZARD for small deviations from $\xi = 0$ already. In addition, for any reasonable NLO observable, as well as the total cross section, the difference in ξ_{min} should not lead to any significant deviations. The results of our comparison studies throughout this thesis support this statement.

The absolute-pseudorapidity $|\eta_{j_i}|$ distributions are shown in fig. 4.2, where we observe a shift towards lower values of $|\eta_{j_i}|$ for the two hardest jets at NLO when compared to LO. In fig. 4.2(c), the $|\eta_{j_3}|$-distribution of the third-hardest jet in case of a real radiation event is given. Just like in the case of the p_{T,j_3}-distribution, this plot has LO accuracy. Contrary to the p_{T,j_3}-distribution, scale variations do not seem to be trivial. This, however, is simply because of higher statistical uncertainties per bin in the scale

[8] This divergent behavior is still regulated by a minimal default cut on the energy of the emitted parton as explained in the following text.

variations themselves.

Similarly to the p_{T,j_3}-distribution earlier, this plot only serves to highlight the differences in the choice of a numerically motivated cutoff ξ_{min} for the lower bound of the energy of the radiated parton. Here, this difference manifests itself in almost insignificant deviations between the MG5_aMC@NLO and WHIZARD results across lower $|\eta_{j_3}|$-regions when only considering statistical errors. However, the significance of the deviation is expected to vanish when also taking scale variations into account.

Nevertheless, varying values of ξ_{min} can lead to significant differences in the $|\eta_{j_3}|$ distribution (when only considering statistical errors). This is shown in the direct comparison between MG5_aMC@NLO and different values for ξ_{min} for WHIZARD in fig. 4.2(d), where we see a good agreement within statistical errors for $\xi_{\mathrm{min}} = 10^{-6}$, similar to the p_{T,j_3}-distribution. This fact, together with the argument that the arbitrary choice of ξ_{min} should not influence the total cross section or any IR-safe NLO observables (as long as ξ_{min} is chosen small enough), strengthens our confidence in our FKS implementation even further, given that we achieve excellent agreement between all three MC event generators for distributions of fNLO accuracy.

Figures 4.3(a) and 4.3(b) show the energy distribution of the p_{T}-hardest and second-hardest jet. In contrast to LO, where both jets are back-to-back and simply have $E_{j_{1/2}} = \sqrt{s}/2$ (at lepton colliders without QED ISR), at NLO we can see the expected shape with a quickly decreasing height of the curve for lower energies due to the real radiation of the gluon. However, because of the fact that all Born-like subevents will fall into the highest-energy bin, the slope is associated with jet energies in real-radiation events. This is again represented in the trivial scale variations in these bins, as there the order of the strong coupling constant is exclusively $\mathcal{O}(\alpha^2\alpha_S)$. For the energy distributions we refrain from showing the distribution of the third jet, as it shows very similar behavior to its p_{T} counterpart and does not give any new insight.

The invariant-mass $M_{j_1 j_2}$ distribution of the system consisting of the two p_{T}-hardest jets is shown in fig. 4.3(c). The expected, steep slope of the distribution towards lower values of $M_{j_1 j_2}$ is purely due to real gluon radiation, a fact that is again evident in the trivial scale variation across the slope. It has to be noted that the lower edge of the scale variation in the highest bin appears unexpectedly low. This is due to the sum of weights of the $\mu = m_Z/2$ scale variation becoming negative in this bin, as explained in the remarks on scale variations at the end of section 4.3.2 regarding negative-valued bins at fNLO. Similar occurrences can be seen for several other differential distributions throughout this thesis.

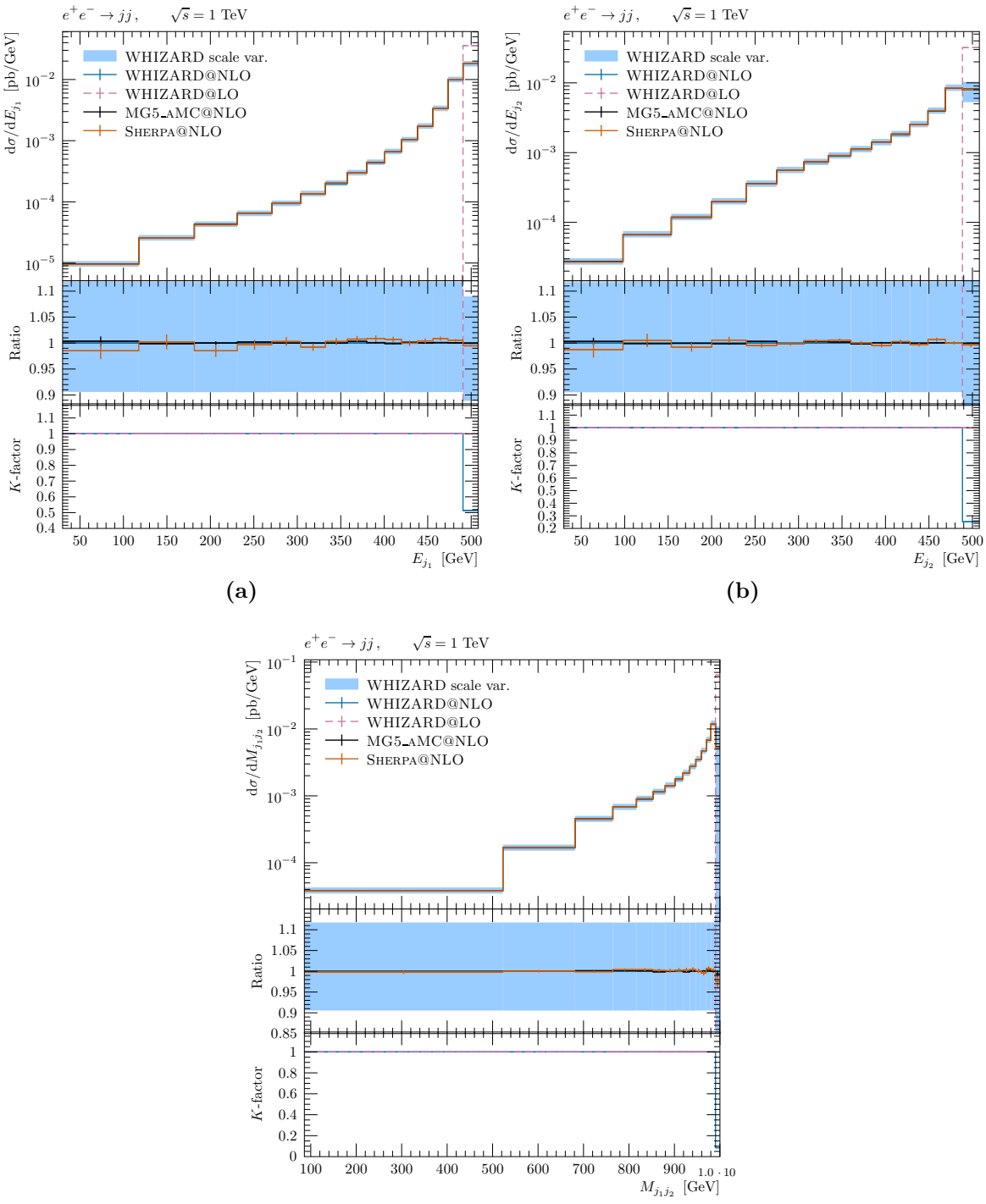


Figure 4.3. Differential distributions of jet energies E_{j_i} for the process $e^+e^- \to jj$ at fNLO QCD of the first- (a), and second-hardest (b) jets, along with (c) the invariant mass $M_{j_1j_2}$ of the system of the two hardest jets. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

Differential Distributions for $e^+e^- \to jjj$

Increasing final-state jet multiplicity to three well-separated jets with the possibility of an additional jet due to real radiation also increases the complexity at NLO, as already described in more detail in the discussion of the fNLO cross sections in section 4.2.2, and in section 3.1.3 regarding the FKS α-regions `alr=5,6` in table 3.1. In short, spin correlations due to final-state $g \to q\bar{q}/gg$ splittings are introduced, and care has to be taken regarding symmetry factors for identical final-state particles and the correct assignment of FKS pairs for α-regions with equal real, but different Born flavor structures. The process $e^+e^- \to jjj$ at fNLO thus provides a good opportunity to further test our FKS implementation in the context of fNLO event generation.

In fig. 4.4, the differential distributions of the transverse momenta p_{T,j_i} of the three p_T-hardest jets are shown. We can see a clear suppression of high-p_T jets in both the distributions for the p_T-hardest and second hardest jets due to the radiation of an additional jet already at LO. The p_{T,j_3} distribution of the third-hardest jet j_3 shows a typical form for a final-state radiated parton, exhibiting a slope from a peak at the p_T cut of 30 GeV towards higher values of p_T, with a quick drop-off at $p_{\mathrm{T},j_3} \gtrsim 170\,\mathrm{GeV}$ as this region becomes increasingly constrained by kinematics.

At NLO, the emitted real parton takes away additional transverse momentum, leading to further suppression of varying degrees across the full p_T spectrum of the three hardest jets, most prominent in higher-p_T regions due to IR-divergent real radiation having dominantly low p_T. This naturally shifts the position of the peaks of the j_1 and j_2 distributions even further to lower p_T than in the LO case, with the p_T of the second-hardest jet peaking as low as approximately 250 GeV. Additionally, again we can see an increasing suppression towards lower values of p_T because of soft radiation making up the majority of low-p_T jets in this phase-space region.

The differential distributions of the jet energies E_{j_i} are shown in fig. 4.5 for the three p_T-hardest jets. Both the E_{j_1} and E_{j_2} distributions show the expected slope from higher to lower energies, whereas the E_{j_3} distribution shows the typical behavior for a radiation from final-state particles, with a slope from lower to higher energies.

The distributions for the two hardest jets both show a suppression in the highest bin because of the energy carried away by the third jet and the real radiation, as both have predominantly low energies. Especially in the case of real emissions in the soft limit, most counterevents, which have 3-parton final states by construction, will decrease the overall sum of weights in the highest bins of the $E_{j_{1/2}}$ distributions. In general, compared to LO, the radiation of an additional parton shifts the energy distribution of j_1 and j_2 significantly to lower energies, while the distribution of the third-hardest jet j_3 experiences lower values across the whole energy spectrum.

In fig. 4.5(d), the invariant mass $M_{j_1 j_2}$ of the system of the two hardest jets is shown.

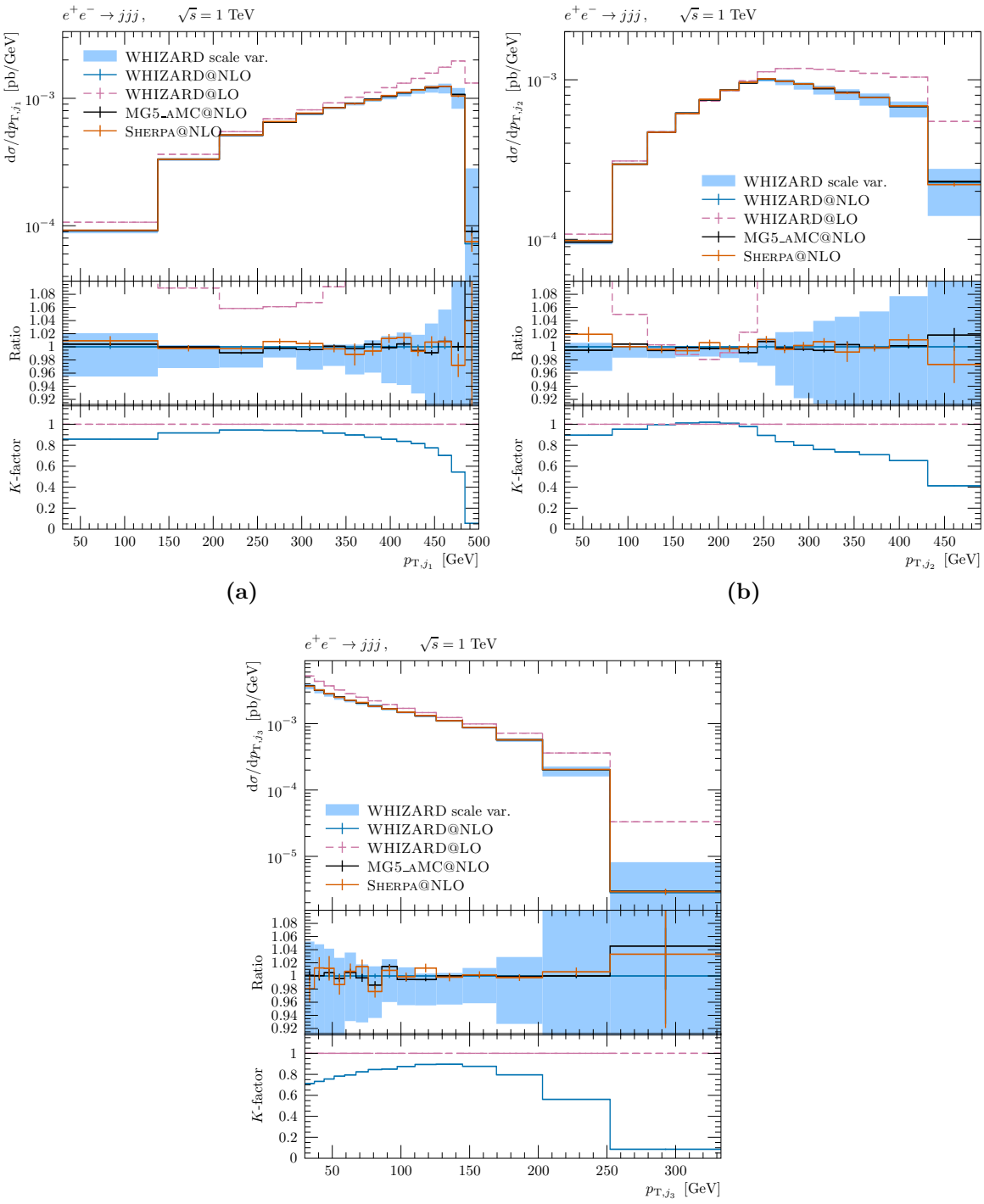


Figure 4.4. Differential distributions of jet transverse momenta p_{T,j_i} for the process $e^+e^- \to jjj$ at fNLO QCD of the first- (a), second- (b), and third-hardest (c) jet. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

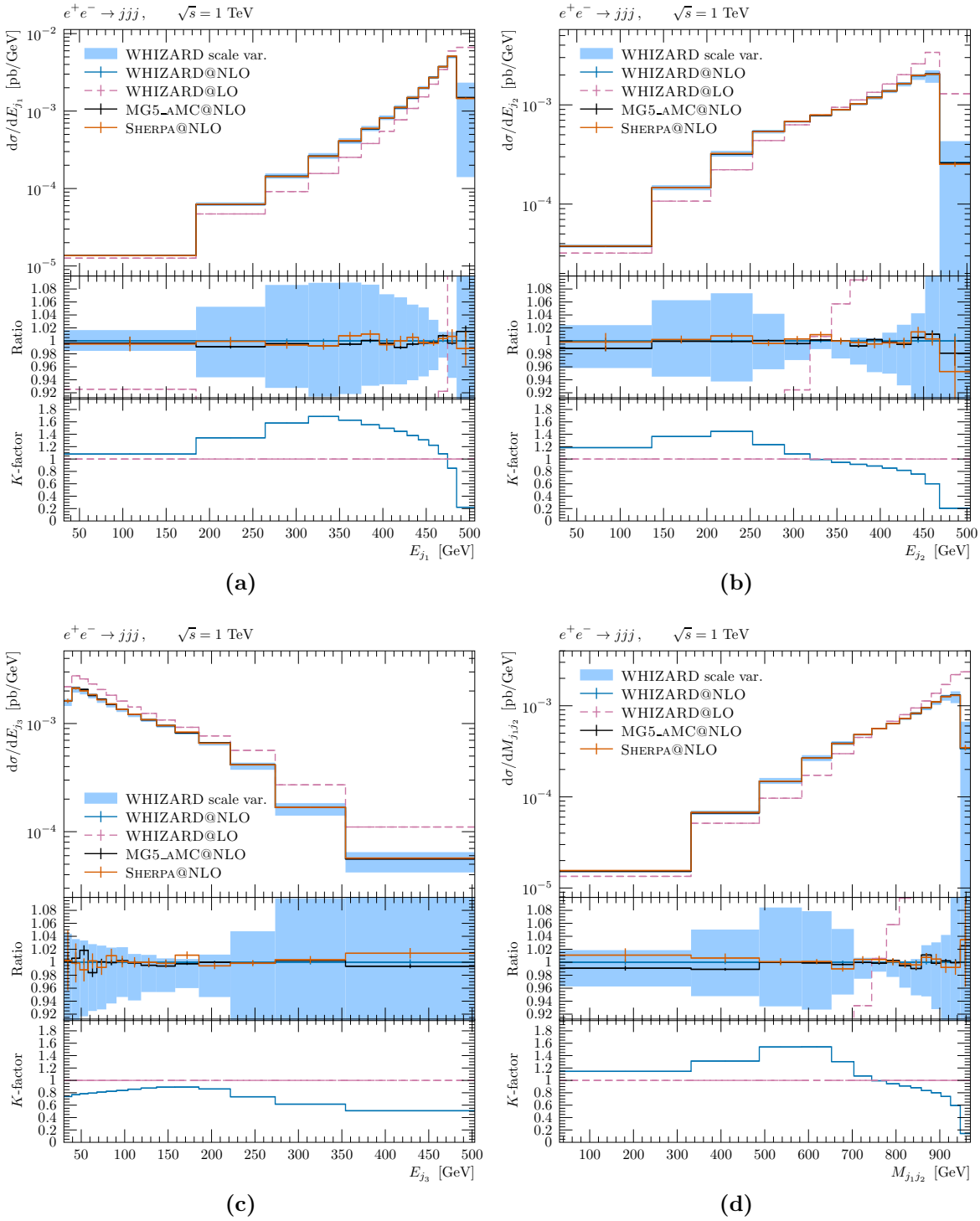


Figure 4.5. Differential distributions of jet energies E_{j_i} for the process $e^+e^- \to jjj$ at fNLO QCD of the first- (a), second- (b), and third-hardest (c) jets, along with (d) the invariant mass $M_{j_1j_2}$ of the system of the two hardest jets. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

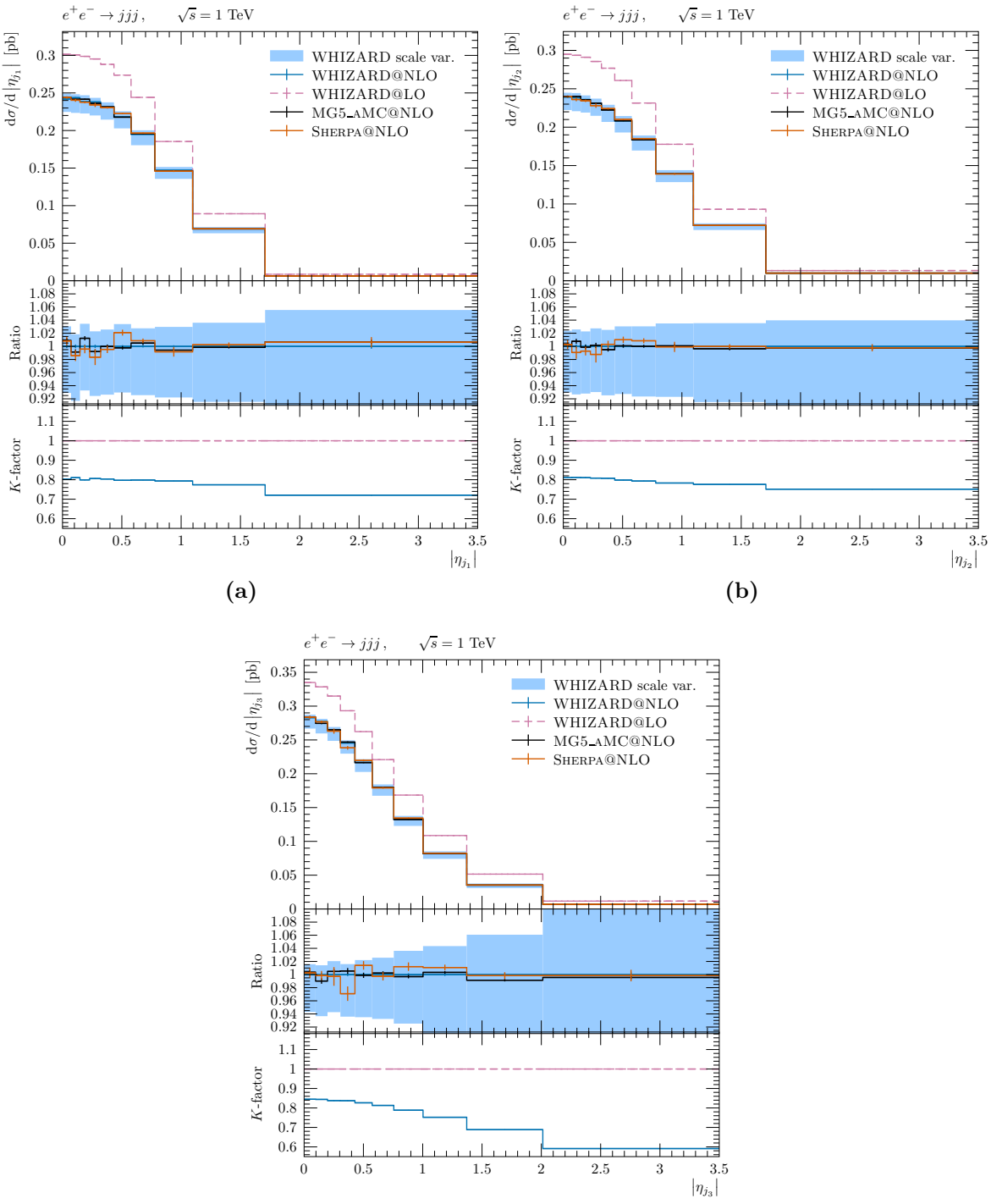


Figure 4.6. Differential distributions of jet pseudorapidities $|\eta_{j_i}|$ for the process $e^+e^- \to jjj$ at fNLO QCD of the first- (a), second- (b), and third-hardest (c) jets are shown. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

It predictably shows similar behavior to the $E_{j_{1/2}}$ distributions, with a suppression of the peak and an overall shift to lower values of $M_{j_1 j_2}$ when comparing the distribution at NLO to the LO case.

In fig. 4.6, the differential distributions of the absolute pseudorapidities $|\eta_{j_i}|$ for the three hardest jets are given. Both the first- and second-hardest jets show a very similar distribution shape, with contributions close to zero above $|\eta_{j_i}| > 1.7$. For the third-hardest jet, the tail of the distribution is stretched out to slightly higher values of $|\eta_{j_i}|$, which is expected, as lower jet-p_T can lead to smaller polar angles. Compared to the LO results, we observe an overall suppression across the whole $|\eta_{j_i}|$ spectrum. In general, we can see a significantly higher negative NLO correction for higher values of $|\eta_{j_i}|$.

Differential Distributions for $e^+e^- \to jjjj$

By increasing the final-state jet multiplicity further to four well-separated jets at Born level, with the possibility of an additional one due to real radiation, we do not introduce new aspects of the FKS subtraction scheme. However, the complexity of the final state increases, leading e.g. to a higher number of α-regions, as well as more complicated spin and color correlations. This puts an excellent stress test on our FKS implementation and internal combinatorics, since effects of small errors might get amplified the higher the final-state multiplicity becomes.

The transverse momenta p_{T,j_i} of the four p_T-hardest jets are shown in fig. 4.7. When comparing the overall shapes of the two hardest jets to the trijet case in fig. 4.4, already at LO we clearly see a shift of the peaks to lower p_T, simply because of the additional radiation of the fourth jet. This obviously also changes the distribution of the third-hardest jet, which experiences a suppression at LO at low jet-p_T due to the fourth-hardest jet j_4 predominantly occupying this phase-space region. Because of the IR-divergent real radiation, we see a general suppression in the higher-p_T regions throughout all four p_T distributions in fig. 4.7 at NLO, as expected.

In the first bin of each p_T histogram of the three hardest jets, the MG5_aMC@NLO results slightly deviate from WHIZARD in terms of the statistical errors, which is likely to be a statistical outlier by MG5_aMC@NLO. This assumption is supported by the excellent agreement between Sherpa and WHIZARD in these bins, especially when we recall that Sherpa uses the CS subtraction scheme instead of FKS and consequently follows a completely different ansatz for the construction of the phase space.

The energy distributions are shown in fig. 4.8. Obviously, they experience very similar effects as already described for the jet-p_T distributions, both for the comparison between three and four jets at LO and between LO and NLO. Here, we also see slight

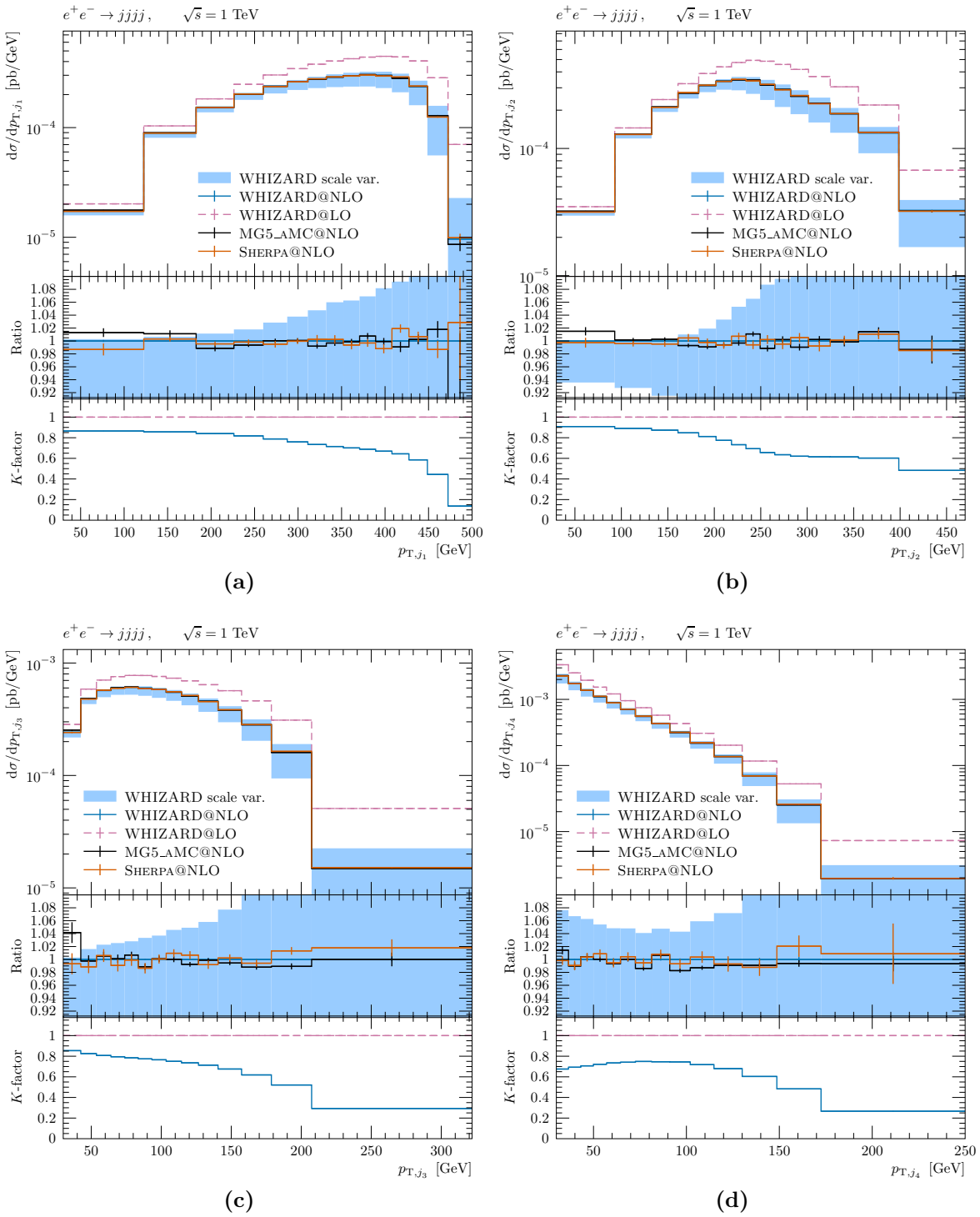


Figure 4.7. Differential distributions of jet transverse momenta p_{T,j_i} for the process $e^+e^- \to jjjj$ at fNLO QCD of the first- (a), second- (b), third- (c), and fourth-hardest (d) jet. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

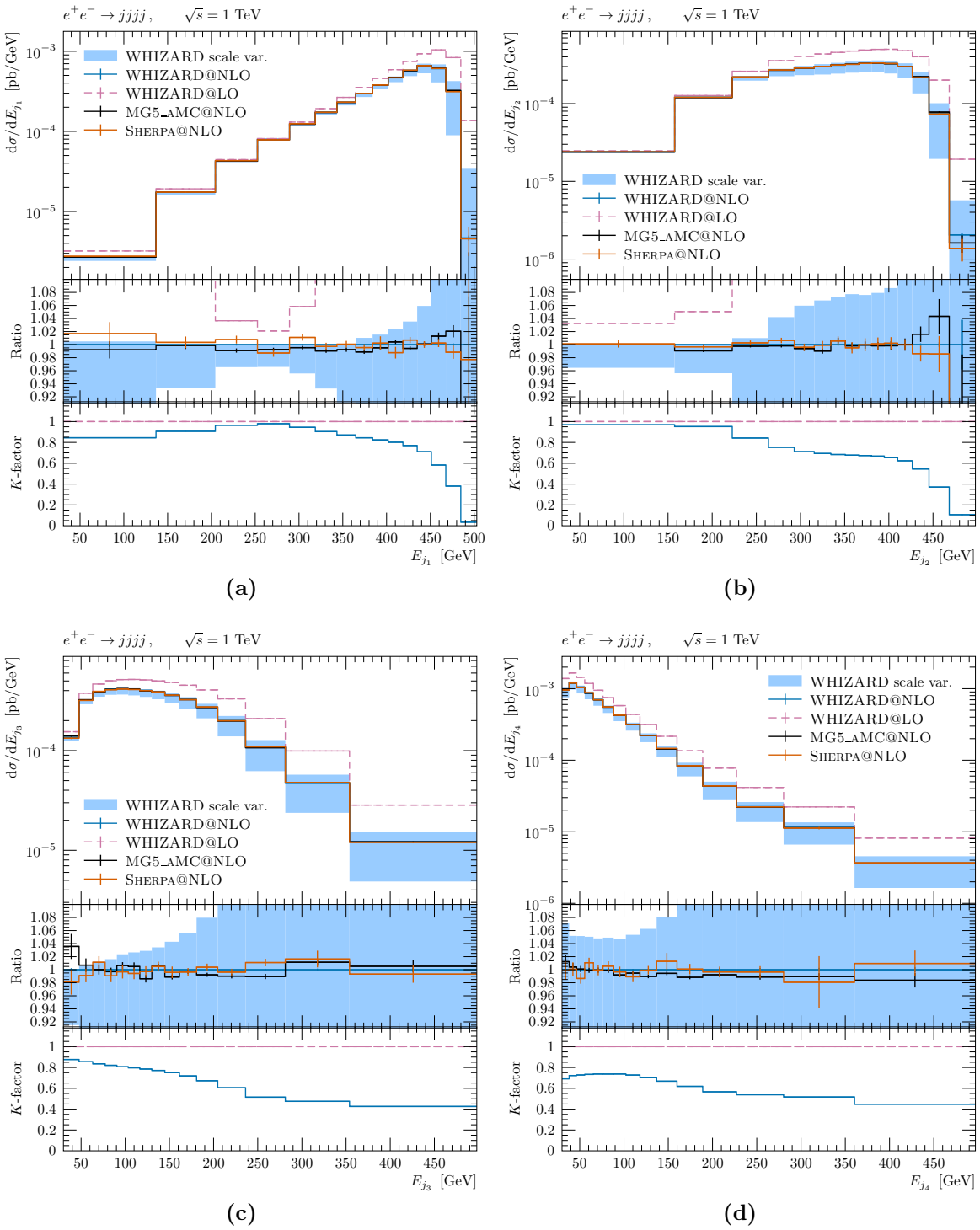


Figure 4.8. Differential distributions of jet energies E_{j_i} for the process $e^+e^- \to jjjj$ at fNLO QCD of the first- (a), second- (b), third- (c), and fourth-hardest (d) jet. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

deviations between the three MC event generators in a few selected bins. These are again nothing more than statistical fluctuations with a significance of $2\sigma^{\text{sig}}_{\text{NLO}}$ to $3\sigma^{\text{sig}}_{\text{NLO}}$ in terms of the statistical errors and do not invalidate our results with WHIZARD. This statement is supported by the fact that in any bin at least one other MC program agrees with WHIZARD well within the statistical and/or theoretical uncertainties.

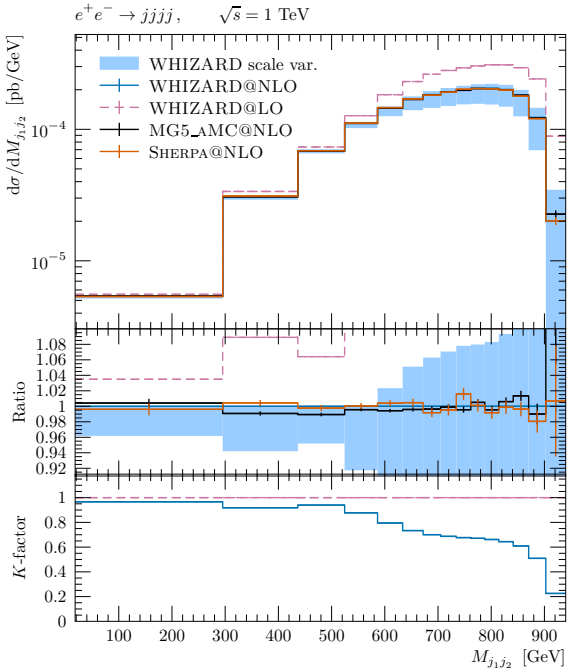


Figure 4.9. Differential distribution of the invariant mass $M_{j_1j_2}$ of the system of the two hardest jets for the process $e^+e^- \to jjjj$ at fNLO QCD. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

The differential distribution of the invariant mass $M_{j_1j_2}$ of the system of the two hardest jets is shown in fig. 4.9. Unsurprisingly, it shows a very similar behavior to the p_{T} and E distributions of the two hardest jets, namely that its peak is shifted to lower values of $M_{j_1j_2}$, both for the comparison between three and four jets at LO and between LO and NLO, and that it shows a stronger suppression for higher values of $M_{j_1j_2}$ due to the real radiation at NLO.

In fig. 4.10, the differential distributions of the absolute pseudorapidities $|\eta_{j_i}|$ for the four hardest jets are given. Similar to the $|\eta_{j_i}|$ distributions for the processes $e^+e^- \to jj$ and $e^+e^- \to jjj$, we see the typical peak at $|\eta_{j_i}|$ close to zero, with a slope towards higher values. This slope is stretched out to slightly higher values of $|\eta_{j_i}|$ the lower j_i is ranked according to its p_{T}, as lower jet-p_{T} can mean smaller polar angles. Again, due to the real radiation being effectively isotropic in terms of the polar angle,

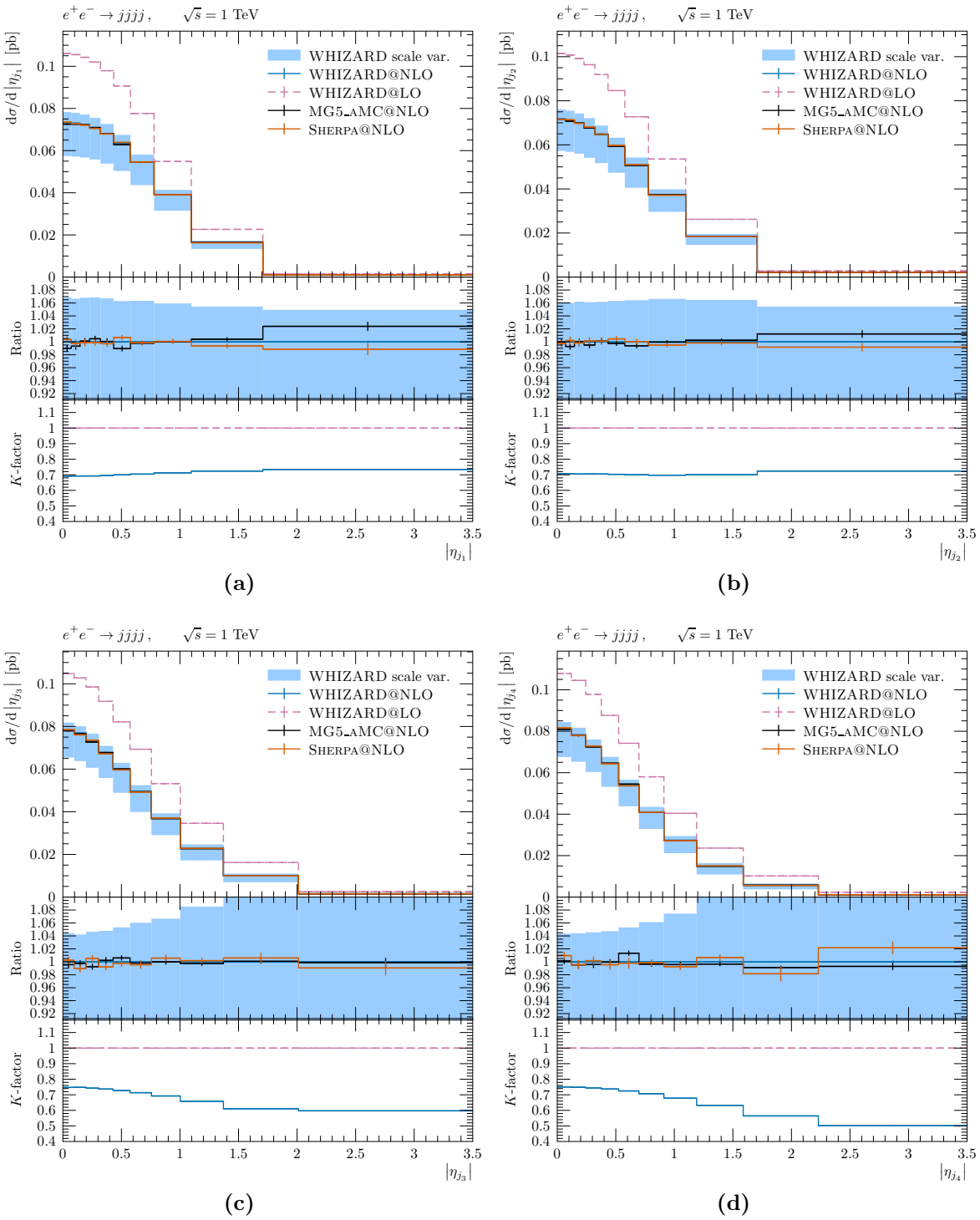


Figure 4.10. Differential distributions of jet pseudorapidities $|\eta_{j_i}|$ for the process $e^+e^- \to jjjj$ at fNLO QCD of the first- (a), second- (b), third- (c), and fourth-hardest (d) jet. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

we observe a mostly uniform suppression across all $|\eta_{j_i}|$ values for j_1 and j_2, while j_3 and j_4 experience a decreasing K-factor towards higher $|\eta_{j_i}|$.

Differential Distributions for $e^+e^- \to jjjjj$

The maximum jet multiplicity for which we are discussing differential distributions is limited by the constraint for a reasonable amount of computing time and precision. Within the time scope of this thesis, these constraints allowed for the generation of FO events for up to five well-separated jets at Born level with the possibility of an additional jet due to real radiation. Yet again, no new aspects of FKS subtraction is introduced with further increasing the jet multiplicity. This also means that, in principle, much higher jet multiplicities are supported by WHIZARD as well, given enough computational capacities. However, as in the jump from three to four jets at Born level, the jump from four to five jets leads to a considerable increase in the number of α-regions. In turn, this introduces even more variation and increases complexity of combinatorics in terms of our FKS subtraction implementation in WHIZARD.

The differential distributions of the transverse momenta p_{T,j_i} of the five p_T-hardest jets are shown in figs. 4.11 and 4.14(a). The trend we have seen so far for increasing jet multiplicity continues here as well, in the sense that the jets show an increasing suppression of higher-p_T regions the lower the p_T ranking of the jet becomes due to additional final-state emission already at LO. This obviously also shifts the peak of the distributions further to the left, with the distribution of the fifth-hardest jet simply exhibiting a slope down from the p_T cut at 30 GeV. At NLO, due to the IR-divergent nature of the real radiation, we see a significantly increasing suppression with higher p_T throughout all jet-p_T histograms.

While Sherpa in general closely agrees with WHIZARD within both the statistical and theoretical uncertainties estimated by scale variations, for many observables that are presented here and in section 4.3.4 MG5_aMC@NLO seems to deviate in some bins from WHIZARD with $\sigma^{\mathrm{sig}}_{\mathrm{NLO}} > 5$ when considering only statistical errors while staying well within the scale variation band of WHIZARD. For example, in fig. 4.11(d) the second-highest bin shows a significance of the deviation between MG5_aMC@NLO and WHIZARD of $7.4\sigma^{\mathrm{sig}}_{\mathrm{NLO}}$. Unfortunately, due to very long run times of the simulation with MG5_aMC@NLO and to an untimely operating system update, we could not conclusively determine whether this is purely a technical error, an inconsistency in our setup of the event generation with MG5_aMC@NLO and the subsequent analysis, or an actual difference between WHIZARD and MG5_aMC@NLO, e.g. in the phase-space treatment for processes with higher multiplicities. However, with the majority of the deviations of MG5_aMC@NLO from WHIZARD still well within the scale uncertainties, as well as with the lack of any significant disagreement between Sherpa

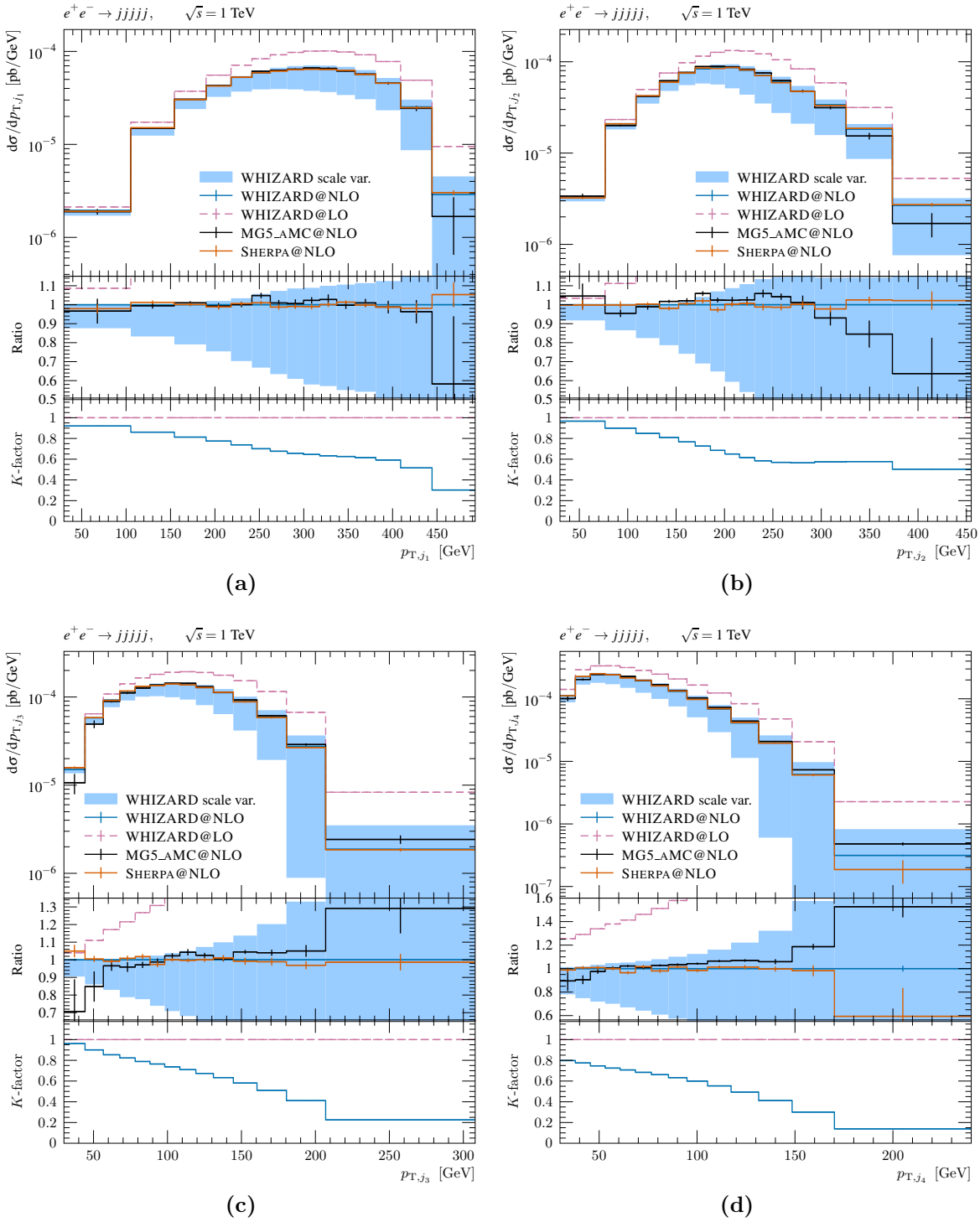


Figure 4.11. Differential distributions of jet transverse momenta p_{T,j_i} for the process $e^+e^- \to jjjjj$ at fNLO QCD of the first- (a), second- (b), third- (c), and fourth-hardest (d) jet. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

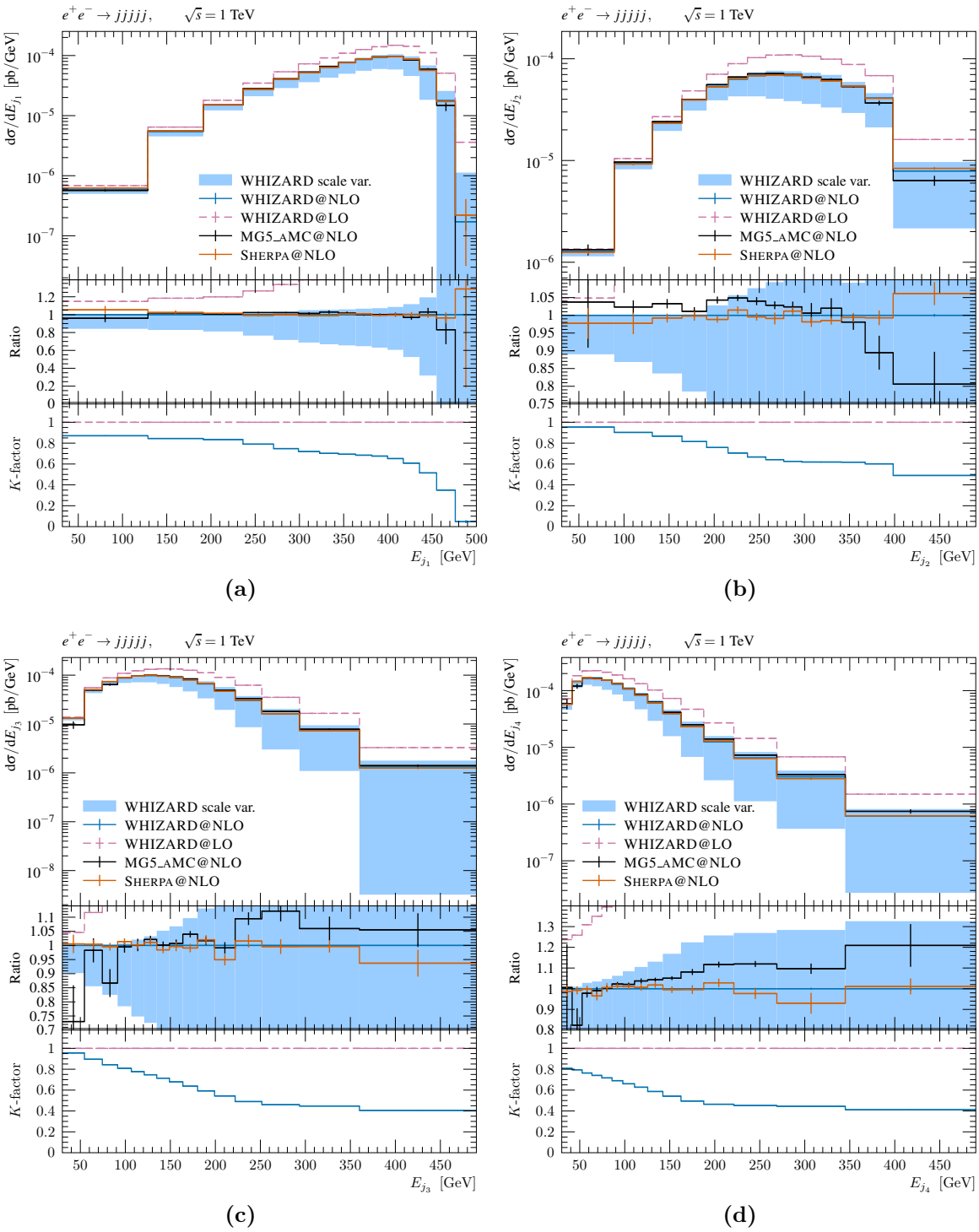


Figure 4.12. Differential distributions of jet energies E_{j_i} for the process $e^+e^- \to jjjjj$ at fNLO QCD of the first- (a), second- (b), third- (c), and fourth-hardest (d) jet. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

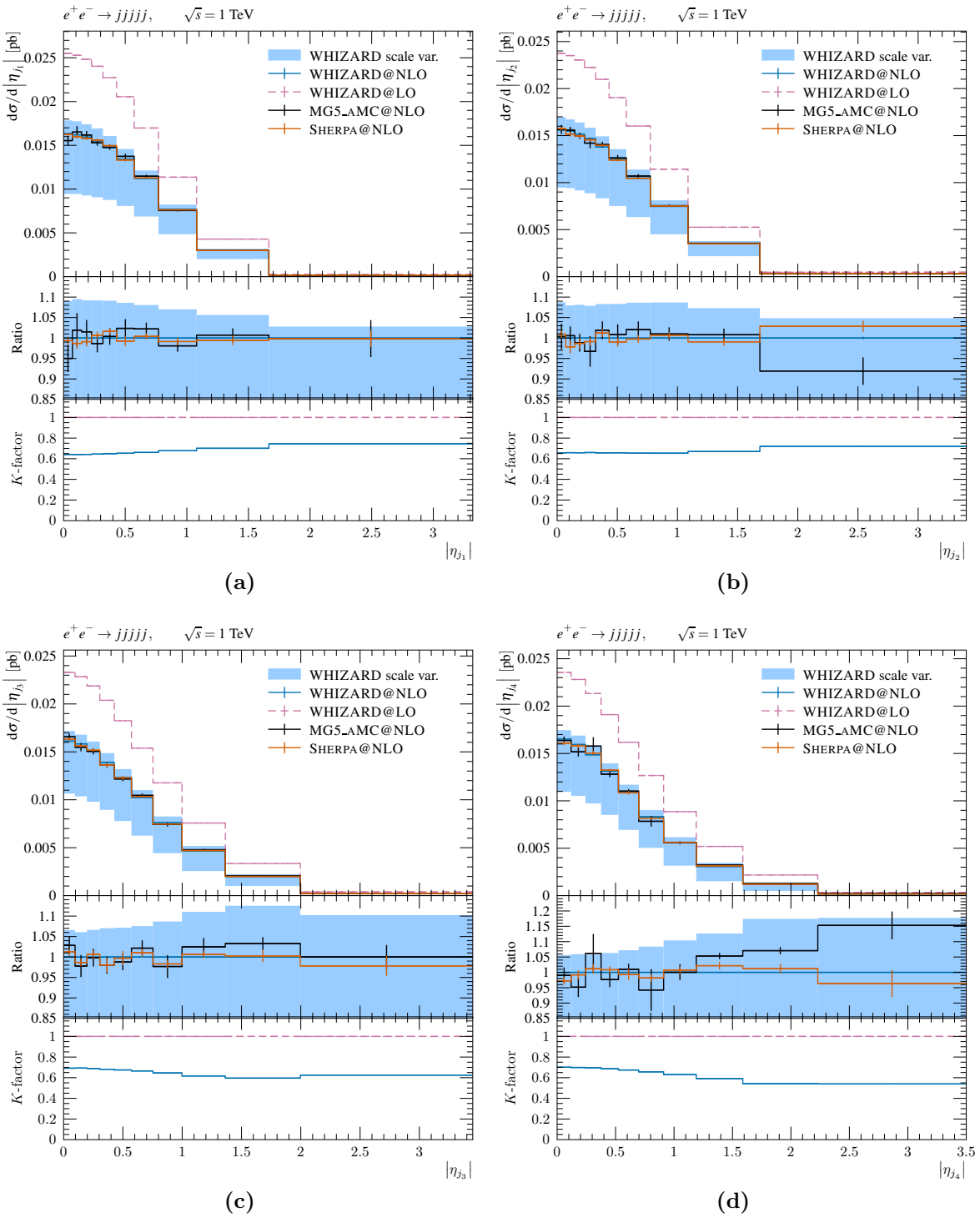


Figure 4.13. Differential distributions of jet pseudorapidities $|\eta_{j_i}|$ for the process $e^+e^- \to jjjjj$ at fNLO QCD of the first- (a), second- (b), third- (c), and fourth-hardest (d) jet. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

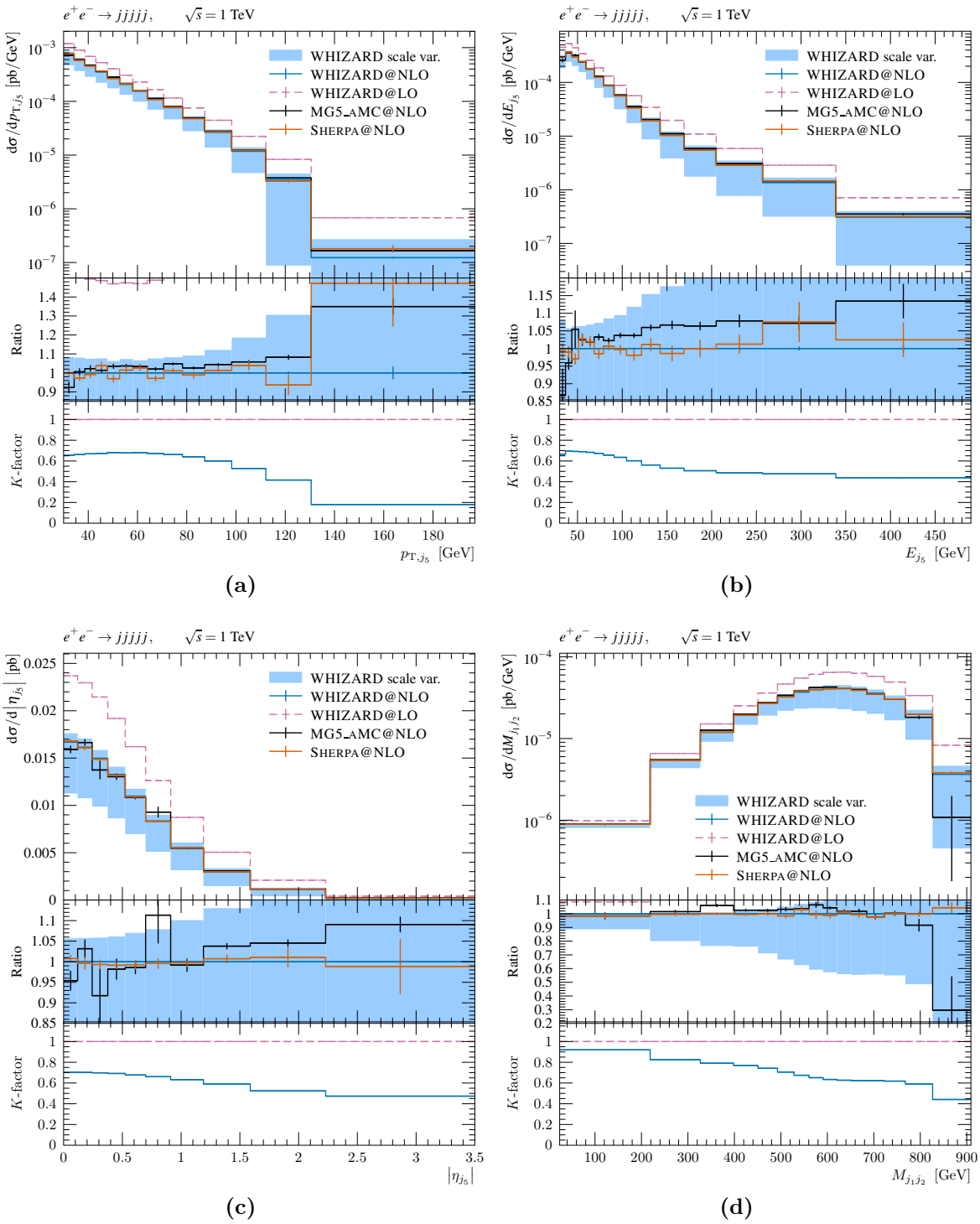


Figure 4.14. Differential distributions of observables (a)-(c) defined on the fifth-hardest jet of the process $e^+e^- \to jjjjj$ at fNLO QCD, along with (d) the invariant mass $M_{j_1j_2}$ of the system of the two hardest jets. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

and WHIZARD, this poses no real concern regarding the validity of our results.

Figures 4.12 and 4.14(b) show the differential distributions of the jet energies E_{j_i}. They obviously behave very similar to the p_{T,j_i} distributions regarding the change from four to five jets at Born level, as well as in terms of the NLO correction due the real emission.

In figs. 4.13 and 4.14(c), the differential distributions of the jet pseudorapidities $|\eta_{j_i}|$ are shown. Again, the behavior is very similar to the trend already seen for lower jet multiplicities. The lower the ranking of the jet-p_{T}, the longer the tail of the $|\eta_{j_i}|$ distribution, as lower p_{T} can lead to smaller polar angles with respect to the beam axis. At NLO, we observe the same behavior as before, namely that the K-factor is roughly constant in each bin for the two hardest jets, with higher negative corrections in the higher $|\eta_{j_i}|$ regions the lower the jet-p_{T} ranking becomes.

Finally, fig. 4.14(d) presents the differential distributions for the invariant mass $M_{j_1 j_2}$ of the system consisting of the two p_{T}-hardest jets. Compared to the 4-jet final state, the high-p_{T} and high-energy regions are more suppressed due to the additional jet, leading to a shift of the $M_{j_1 j_2}$ peak further to the left. At NLO this is even more pronounced, as the real radiation takes away additional energy and momentum. Since the real radiation is predominantly soft in both p_{T} and E, the highest corrections occur in the high-$M_{j_1 j_2}$ regions.

4.3.4. Event-shape Distributions for $e^+e^- \to n$ Jets

In order to further validate NLO effects of final-state radiation from pure n-jet final states in e^+e^- collisions, we will now focus on the differential distributions of the event-shape observables defined in section 4.3.2. To some extent, comparing event-shape distributions between MC event generators is "purer" than comparing observables defined on jets obtained by a jet algorithm, as differences or errors in the respective interfaces to FastJet could lead to deviations between the differential distributions generated by the programs.

Albeit the fact that we do not observe any truly-significant deviations between MG5_aMC@NLO, Sherpa, and WHIZARD in section 4.3.3, it is still important to validate our results for event-shape observables since they can be important tools both for precision studies and BSM searches, as detailed in section 4.3.2. The behavior of event-shape observables is best visualized by directly comparing histograms for different jet multiplicities, which is why in the following we present results grouped by means of observable and not by the jet multiplicity of the process as in section 4.3.3.

Again, to avoid repeating the same observation for the majority of histograms that are shown and discussed in the following, we restate the same preemptive conclusion we draw at the beginning of section 4.3.3: In general, we see a very good agreement between

MG5_aMC@NLO, Sherpa, and WHIZARD, with deviations predominantly below the order of a few percent from the central value of WHIZARD and well within the theoretical uncertainties estimated by scale variations. In addition, the majority of these deviations are within a significance of $1\sigma^{\text{sig}}_{\text{NLO}}$ to $2\sigma^{\text{sig}}_{\text{NLO}}$ in each bin when only considering statistical errors. If necessary, we will discuss any outliers individually.

The differential distributions for the thrust $1 - T$ are shown in fig. 4.15 for all jet multiplicities $n_j = 2, \cdots, 5$. Exactly as expected, in the LO dijet case we see a single spike at zero, as all events have two perfectly back-to-back jets in the CM system. At NLO, the peak at zero becomes significantly suppressed and we see a decreasing slope towards higher values of $1 - T$ due to real-radiation events with values up to roughly $1 - T = 0.35$. This is perfectly in line with the expectation of typical values of $1 - T \leq 1/3$ for trijet events.

The shape of the thrust distribution in the dijet case at NLO is closely repeated in the LO trijet case, with a hard cut-off at similar values. Because of our IR-safe jet-selection criteria (in this case at least three well-separated jets), a pencil-like event shape becomes impossible, excluding values very close and equal to zero. At NLO, again, lower thrust values are suppressed by the real emission and the tail of the distribution reaches out to values above the maximum observed for LO, as the events become more spherical in shape due to the additional radiation. This trend continues for higher jet multiplicities.

However, while the peak of the distributions experience visible shifts to higher $1 - T$ values as the events become more and more spherical with higher n_j, the maximum value does not increase much between the plots for the two highest n_j. This is in line with our expectation, as the upper bound is $1 - T = 1/2$. Moreover, the shift of the peak towards higher values becomes less pronounced for $n_j + 1$ the higher n_j already is. The same can be observed for the amplitude of the K-factor when going from four to five jets at Born level. To some degree, this overall behavior decreases the ability of the thrust to differentiate between events with different, but high jet multiplicities. This can be seen for some other event shapes throughout this section to some extent as well.

Closely related to thrust is the thrust major T_{M}. Its distribution for each jet multiplicity is shown in fig. 4.16. Since T_{M} is defined by the thrust major axis $\mathbf{n}_{T_{\text{M}}}$ that maximizes eq. (4.3.5) in the plane perpendicular to the thrust axis $\mathbf{n}_T$, it is zero for perfectly back-to-back events and becomes larger the more the shape of an event extends into the plane perpendicular to its thrust axis. This is exactly what we observe in the T_{M} distributions. For dijet events at LO we have a single spike at zero, which is significantly suppressed at NLO due to real emissions diverting the event shape away from a pencil-like shape at LO, leading to the formation of a tail towards higher values of T_{M}. The distributions for $n_j > 2$ again cannot reach values very close or equal to

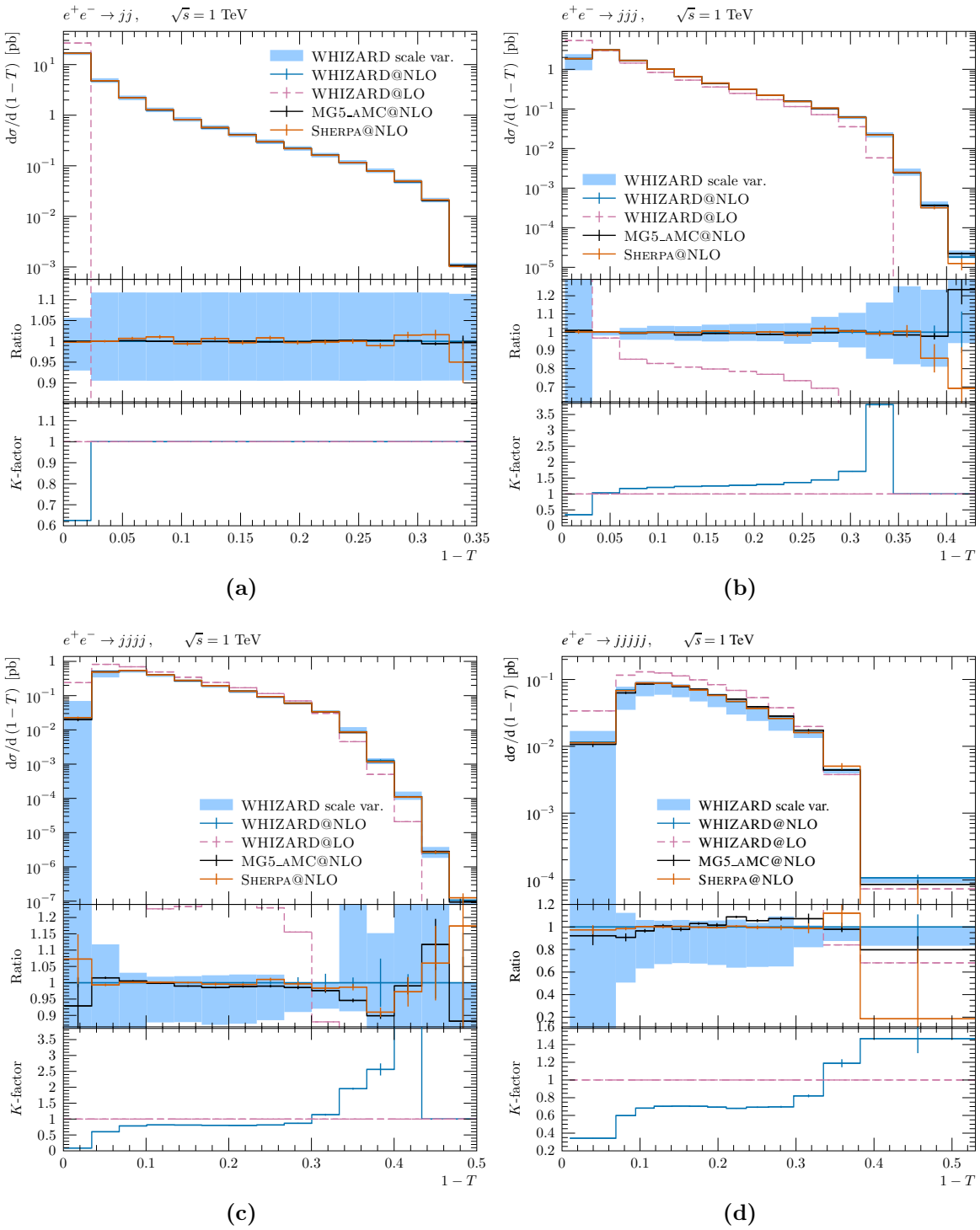


Figure 4.15. Differential distributions of the thrust $1-T$ for different jet multiplicities at fNLO QCD. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result. We note that in (d) the thrust range extends beyond its upper bound of $1-T = 1/2$. This is due to an unfortunate binning.

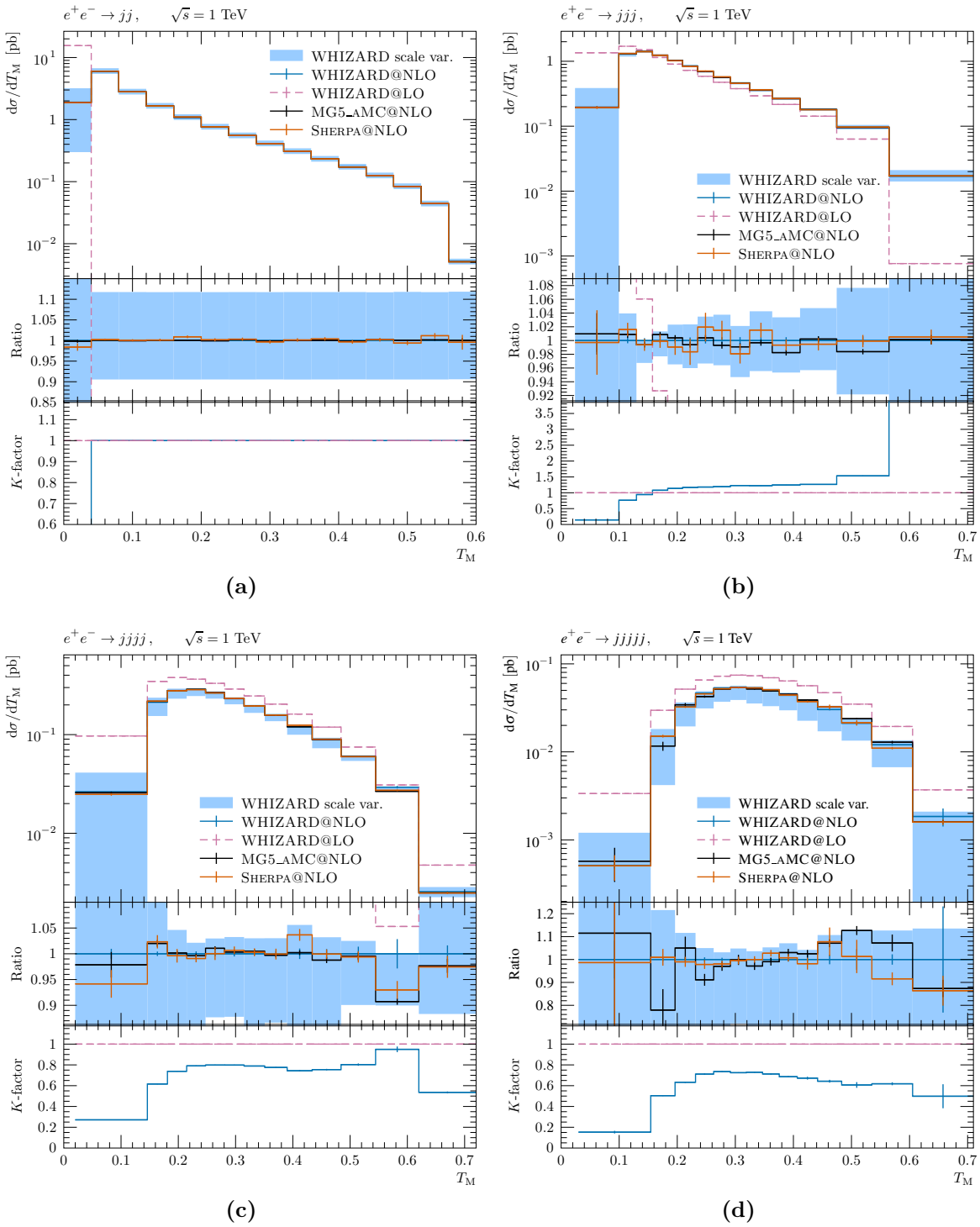


Figure 4.16. Differential distributions of the thrust major T_M for different jet multiplicities at fNLO QCD. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

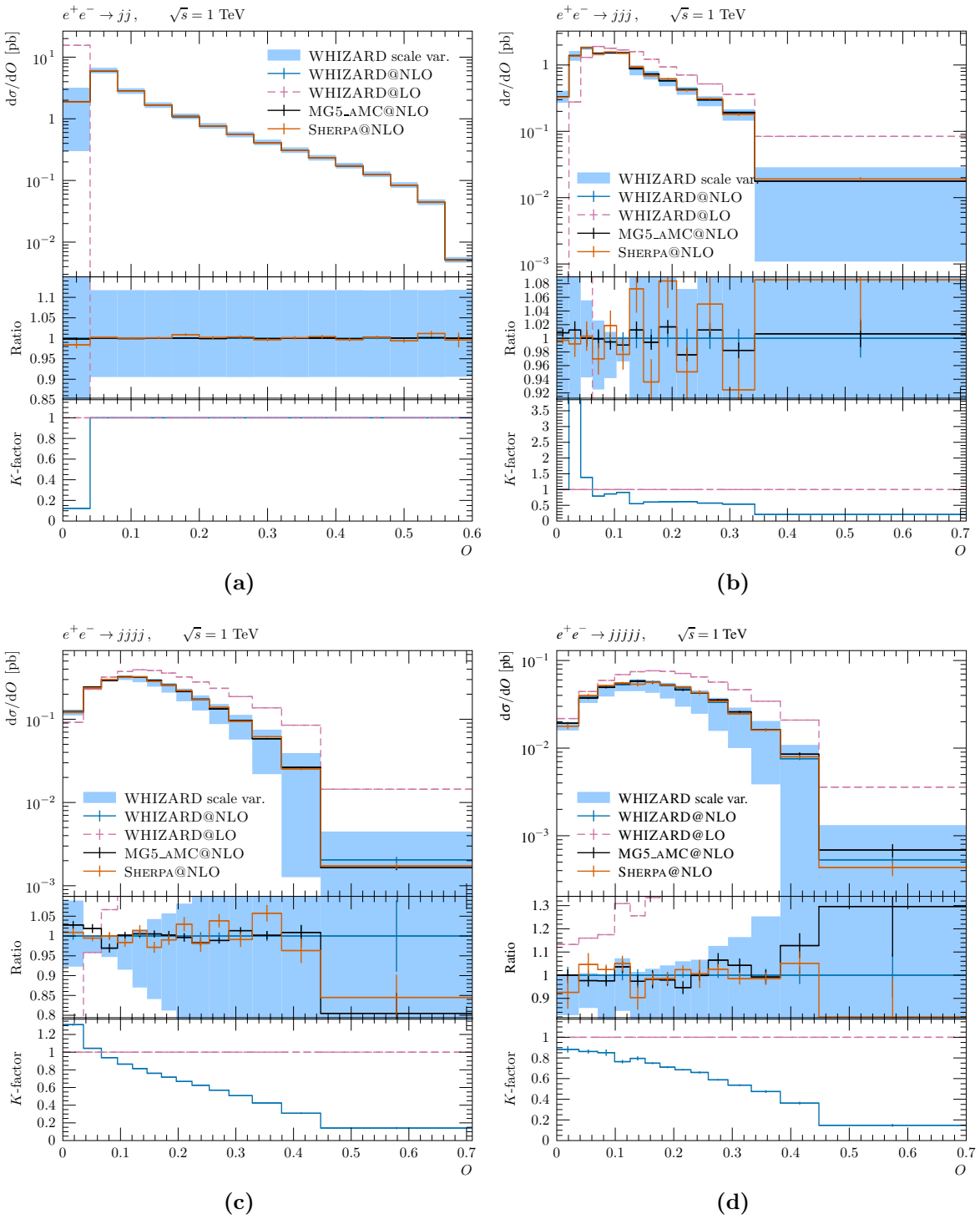


Figure 4.17. Differential distributions of the oblateness O for different jet multiplicities at fNLO QCD. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

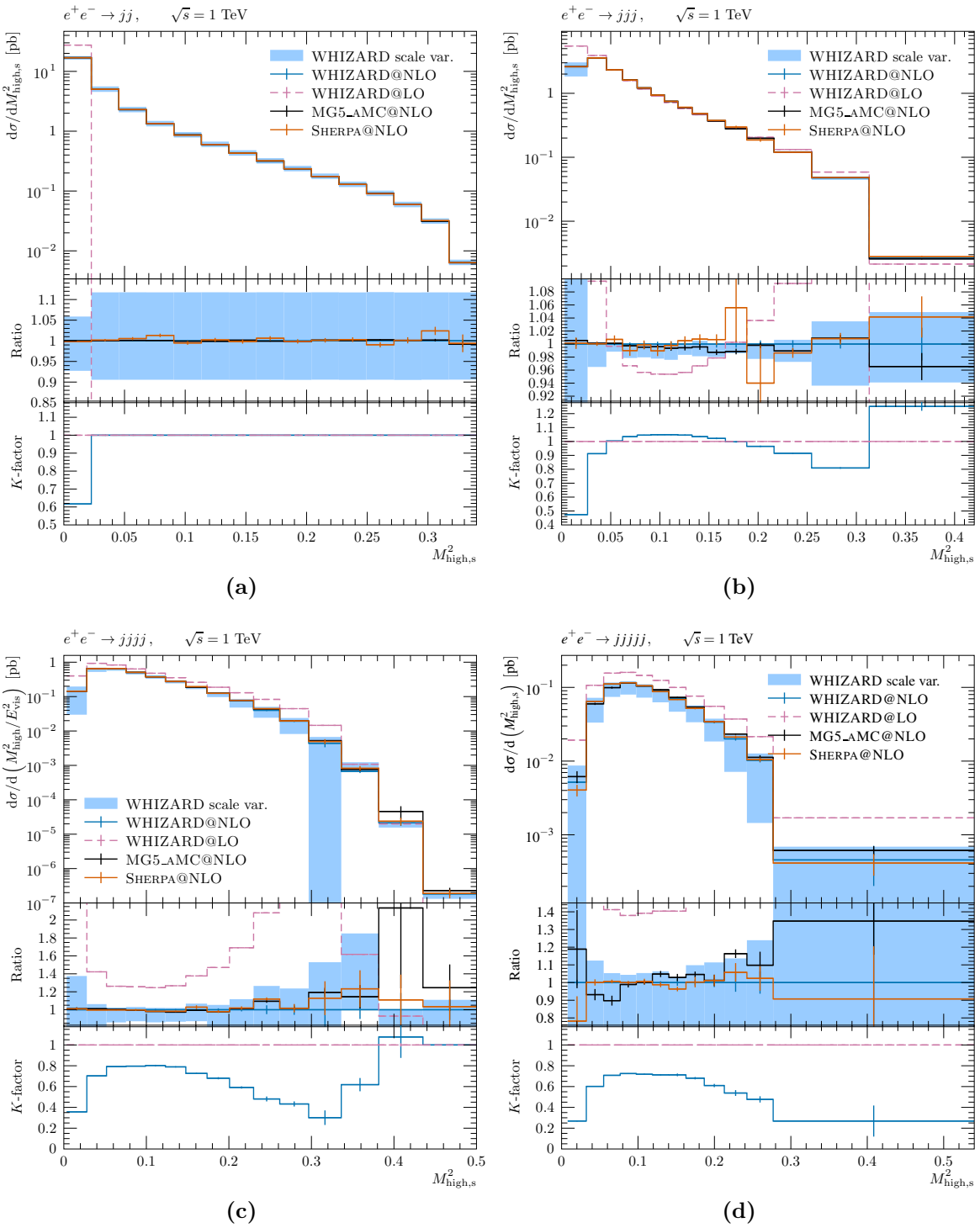


Figure 4.18. Differential distributions of the scaled high hemisphere squared mass $M^2_{\mathrm{high,s}}$ for different jet multiplicities at fNLO QCD. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

zero, simply because these regions are excluded by our jet selection criteria. While there is a significant shift to higher T_{M} values when comparing NLO to LO in the trijet case, NLO corrections for higher n_j are generally negative in all bins. Overall, however, we again observe that the peak of the distribution shifts to higher values for higher n_j, indicating an expansion of the overall event shape in the plane perpendicular to the thrust axis as the event becomes more and more spherical.

The differential distributions for the oblateness O for different jet multiplicities is given in fig. 4.17. It is the difference between thrust major T_{M} and minor T_{m}, and because of the fact that T_{m} becomes different from zero only for jet multiplicities $n_j \geq 4$, the oblateness distribution is equal to $1 - T$ in the dijet case, both at LO and NLO. In the trijet case, we can observe a significant shift of the distribution to lower values of O when comparing NLO to LO, with a strong negative correction the higher the value of O becomes. This can be observed for higher n_j as well and is expected, as typical values for T_{m} become closer to T_{M} the more spherical an event becomes. Thus, typical values for oblateness are maximized for small jet multiplicities higher than $n_j = 2$, while for large enough n_j we expect the peak of the distribution to close in on zero again.

Turning our attention to the hemisphere masses and broadenings, we plotted the scaled high hemisphere squared mass $M^2_{\mathrm{high,s}}$ in fig. 4.18 for the different jet multiplicities. Again, starting with the dijet case, we observe the expected behavior of an increasing suppression of regions close to zero as we move to higher jet multiplicities, both when comparing iterations of n_j at LO and when considering NLO corrections for each n_j. Also, both the extension of the tail and the peak of the distributions move to higher values of $M^2_{\mathrm{high,s}}$ with higher n_j, with typical values still staying well below $M^2_{\mathrm{high,s}} = 0.5$, a value which would indicate a perfectly even distribution of masses in both hemispheres $S_{\pm}$. In the trijet case in fig. 4.18(b), the ratio of Sherpa over WHIZARD shows two bins around $M^2_{\mathrm{high,s}} \approx 1.9$ with large errors and relatively high amplitudes in opposite directions. This is a prime example of the effects of misbinning, discussed in section 3.1.5.

The differential distributions for the wide hemisphere broadening B_{w} are shown in fig. 4.19 for the different jet multiplicities. Overall, the typical behavior for event shapes seen throughout this section can be observed again, with lower values of B_{w} becoming increasingly suppressed for higher jet multiplicities. Also, while the tail of the distribution quickly reaches an maximum around $B_{\mathrm{w}} \approx 0.35$, as the possible increase in maximal broadness is highest for the jump from a dijet to a trijet configuration, the peak of the distribution is shifting visibly to the right as more spherical configurations become kinematically more probable with higher multiplicities.

The total hemisphere broadening B_{tot}, given in fig. 4.20, shows an overall very similar behavior. However, while the NLO correction for B_{w} is dominantly negative, we see a

shift from negative to positive corrections for higher values of B_{tot} when taking into account the sum of broadenings from both hemispheres.

In fig. 4.21, the differential distributions for the Parisi event-shape variable C are given for all jet multiplicities. As C can be understood as a measure for the planarity of an event, it is not surprising to see the strong difference in the distributions for LO and NLO in the dijet case, as the addition of a real emission leads to the typical decreasing slope away from the single spike at zero at LO. However, C shows a strong potential for differentiating jet multiplicities above $n_j = 3$ as well. When looking at the trijet case, a significant extension of the tail towards higher C values can be seen at NLO when compared to LO, with an overall shift of the distribution and its peak to the right. This shift of the peak of the distribution to the right, which can be seen in most other event-shape observables to varying degrees of intensity, continues to be very pronounced even in the 5-jet case. The behavior of the K-factor is similar to that of the total hemisphere broadening, with high negative corrections for low values of C, followed by comparably less negative corrections over a large middle range, with strong positive corrections (or relatively low negative ones in the case of 5-jet) in the higher regions.

The Parisi event-shape variable D, given in fig. 4.22 for the jet multiplicities $n_j = 3, 4, 5$, can be understood as a measure of the sphericity of an event. It is only different from zero for events with more than three jets. Therefore, the distribution for the dijet case is not shown here, as it would be trivial. Instead, the typical behavior for increasing jet multiplicities we have seen for other event-shape observables now starts with a single spike at zero in the trijet case at LO. At NLO, we observe a slope over the full range of D from low to high values. Moving to higher n_j, we then again observe the shift of the peak to the right and an increasing suppression of lower values of D. However, even for the 5-jet process at NLO, most of the events exhibit a very low value for D, indicating a sphericity that is not very pronounced yet.

For $n_j = 5$, we observe a deviation of MG5_aMC@NLO from WHIZARD well outside the scale uncertainty for $D \gtrsim 0.15$. When considering statistical uncertainties, several bins show seemingly significant deviations between MG5_aMC@NLO and WHIZARD of $\sigma^{\text{sig}}_{\text{NLO}} > 5$. However, we would expect both programs to still agree well within the theoretical uncertainties if scale variations were considered with MG5_aMC@NLO as well. Unfortunately, the required run time by MG5_aMC@NLO for this endeavor proved to be prohibitive within the time scope of this thesis. In the future, it should be clarified if this behaviour is due to the same possible reasons stated already in section 4.3.3 in the context of $n_j = 5$.

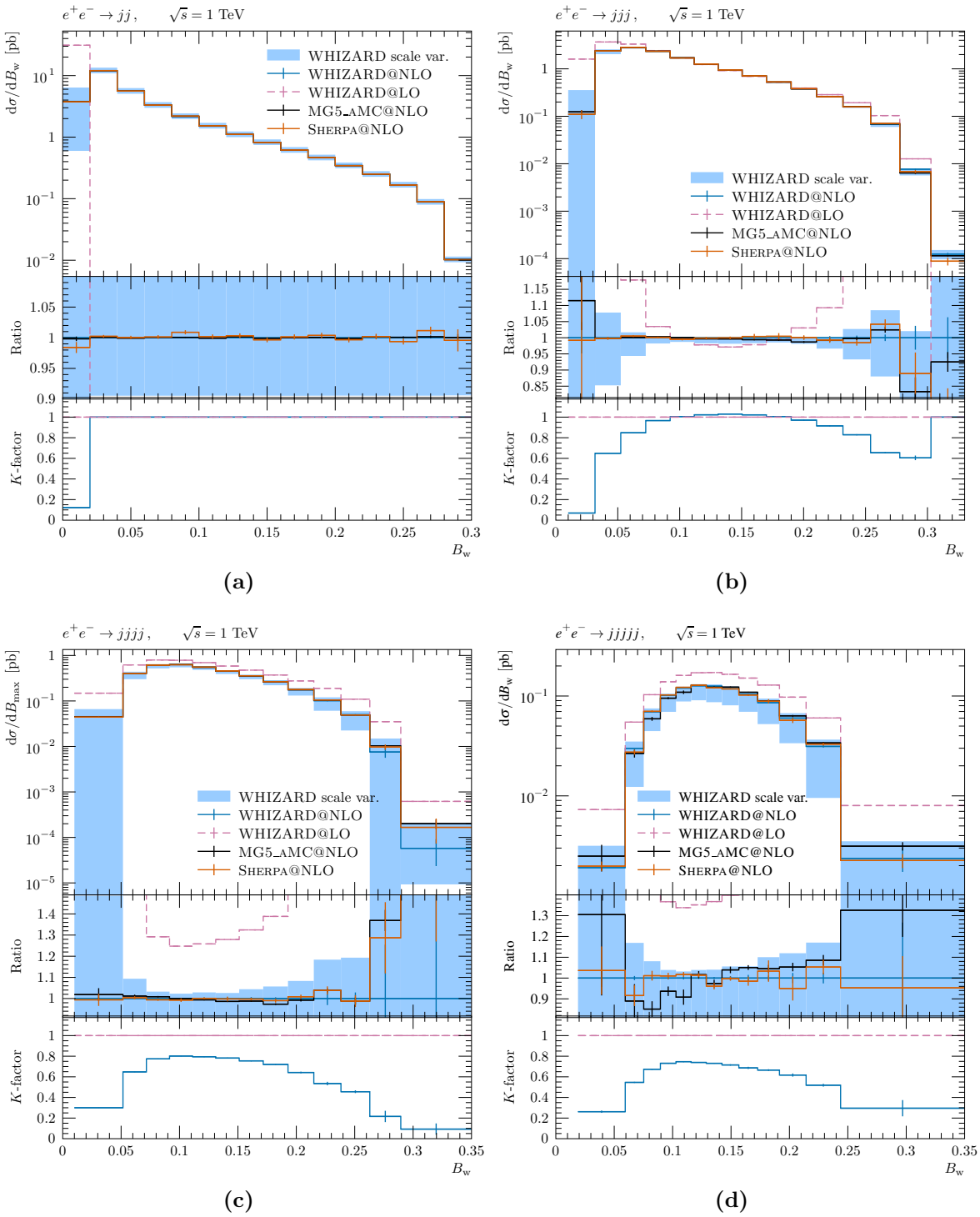


Figure 4.19. Differential distributions of the wide hemisphere broadening B_w for different jet multiplicities at fNLO QCD. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

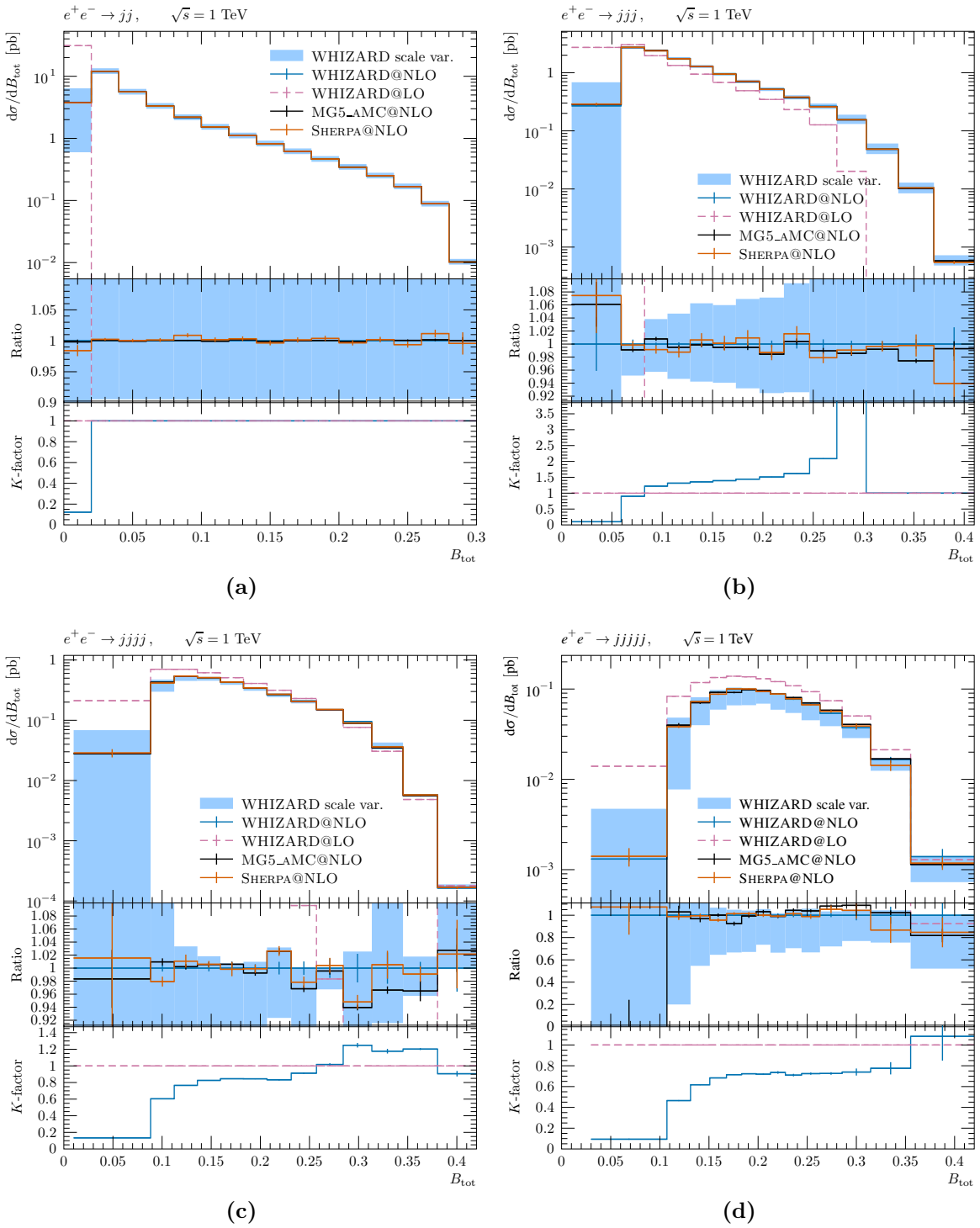


Figure 4.20. Differential distributions of the total hemisphere broadening B_{tot} for different jet multiplicities at fNLO QCD. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

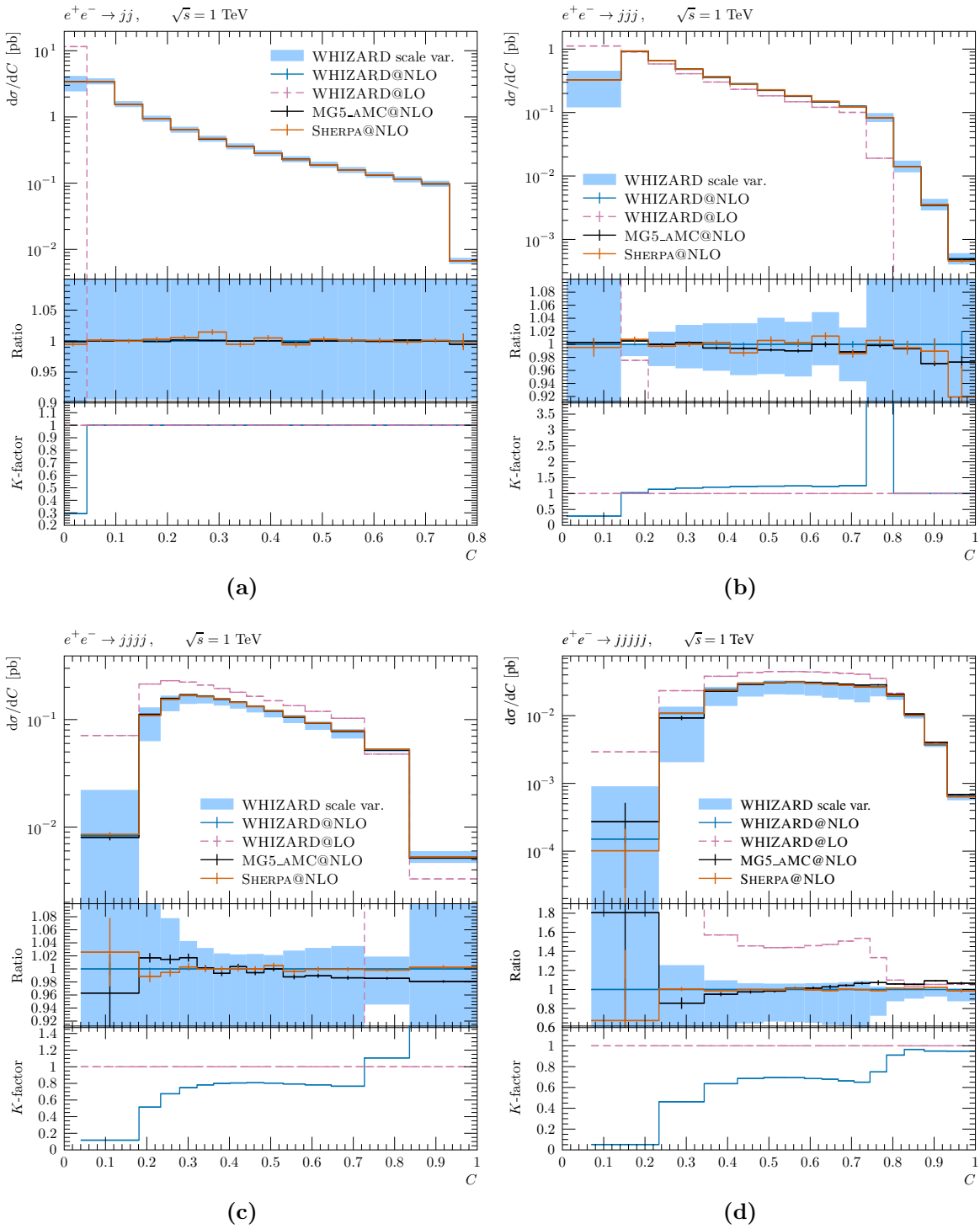
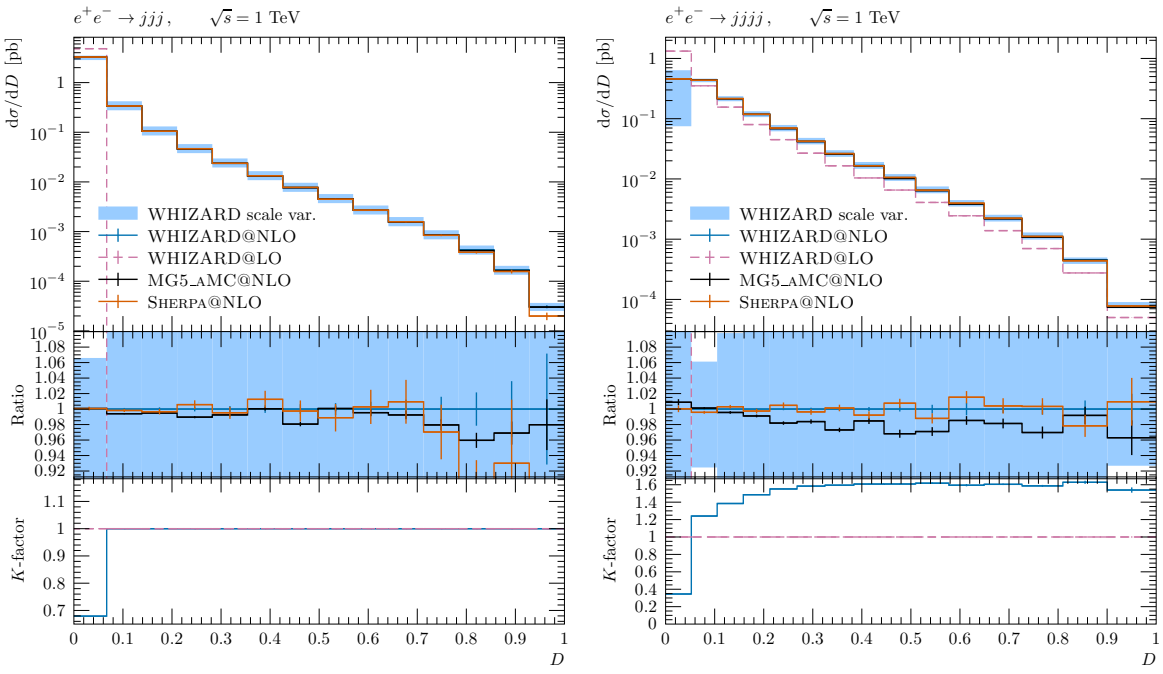
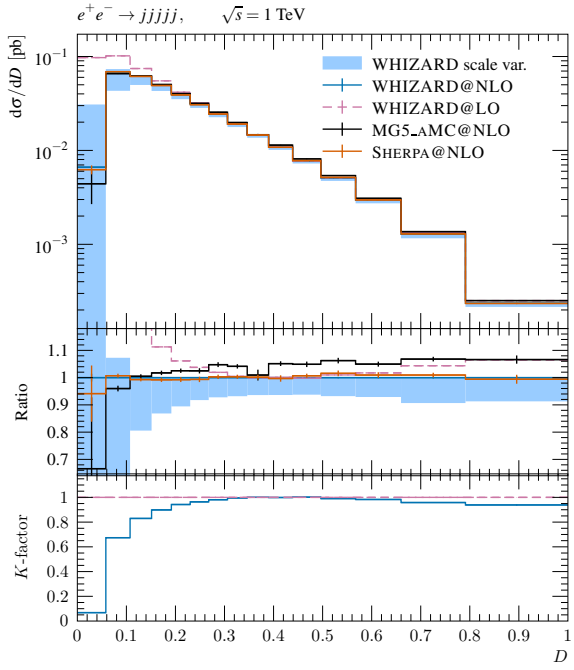


Figure 4.21. Differential distributions of the Parisi event-shape variable C for different jet multiplicities at fNLO QCD. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.



(a) (b)



(d)

Figure 4.22. Differential distributions of the Parisi event-shape variable D for different jet multiplicities at fNLO QCD. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

4.3.5. Differential Distributions for $pp \to Zj$

While we focussed our efforts for the validation of event generation at fNLO QCD mostly on leptonic initial states and thus FSR, we will shortly discuss results for the hadron-collider process $pp \to Zj$ in the following. Here, the system of emitting and emitted parton recoils against the massive Z boson, giving rise to more complex kinematics. This process features both the full range of possible final-state splittings shown in figs. 2.2 and 2.3 due to the jet in the final state at Born level, as well as the initial-state splittings shown in fig. 2.5.

The differential distributions for the transverse momentum $p_{\mathrm{T},Z}$ and p_{T,j_1} of the Z boson and the p_{T}-hardest jet j_1, respectively, as well as their absolute[9] pseudorapidities $|\eta_Z|$ and $|\eta_{j_1}|$ are given in fig. 4.23. All three MC event generators agree well within the theoretical uncertainties estimated by scale variations. Even when only considering statistical uncertainties, the programs agree in almost every bin and any deviations are at the level of a few percent at most.

Looking at the p_{T} distribution of the Z boson in figs. 4.23(a) and 4.23(b), at LO we can see a rising contribution towards lower p_{T} with a hard cut-off below 30 GeV.[10] The reason for the latter is of course the cut on the jet transverse momentum of $p_{\mathrm{T},j_1} > 30\,\mathrm{GeV}$, which indirectly restricts the minimum p_{T} of the Z boson kinematically. At NLO, this restriction is lifted by the presence of an additional jet from the real emission, which can increase or lower the sum of jet-p_{T}. The Z boson distribution at NLO can thus exhibit p_{T} much lower than 30 GeV. We also observe positive NLO corrections across the tail of the distribution towards higher p_{T}.

The transverse momentum distribution of the hardest jet at LO, which originates from an initial-state splitting, shows a decreasing slope towards higher jet-p_{T}, with a peak at the cut-off of $p_{\mathrm{T},j_1} = 30\,\mathrm{GeV}$. This is expected, as the radiation of soft jets is favored. At NLO, we observe positive corrections over the whole p_{T} range, with a significantly increasing K-factor towards higher values of p_{T}.

When examining the pseudorapidity distributions in figs. 4.23(c) and 4.23(d), for the Z boson we see a clear peak slightly above $|\eta_Z| = 2$, meaning most of the Z bosons are produced with a central momentum direction (i.e. at a rather small angle to the beam axis). A reason for this boost along the z-axis is the initial-state emission of a parton already at Born level, which changes the momentum of the emitting parton, leading to differences in the momentum of the two quarks that form the Z boson.

This disparity is enhanced further if the jet at Born level was produced by ISR of a gluon. In this case, the quark can either be a valence or a sea quark and the anti-quark

[9] We checked that the pseudorapidity distribution is symmetric, which allows us to increase statistics by using the absolute value.

[10] Due to the binning being done according to a Breit-Wigner distribution applied to the full range of values at NLO, the lowest bin edge at LO in fig. 4.23(a) is less than 30 GeV.

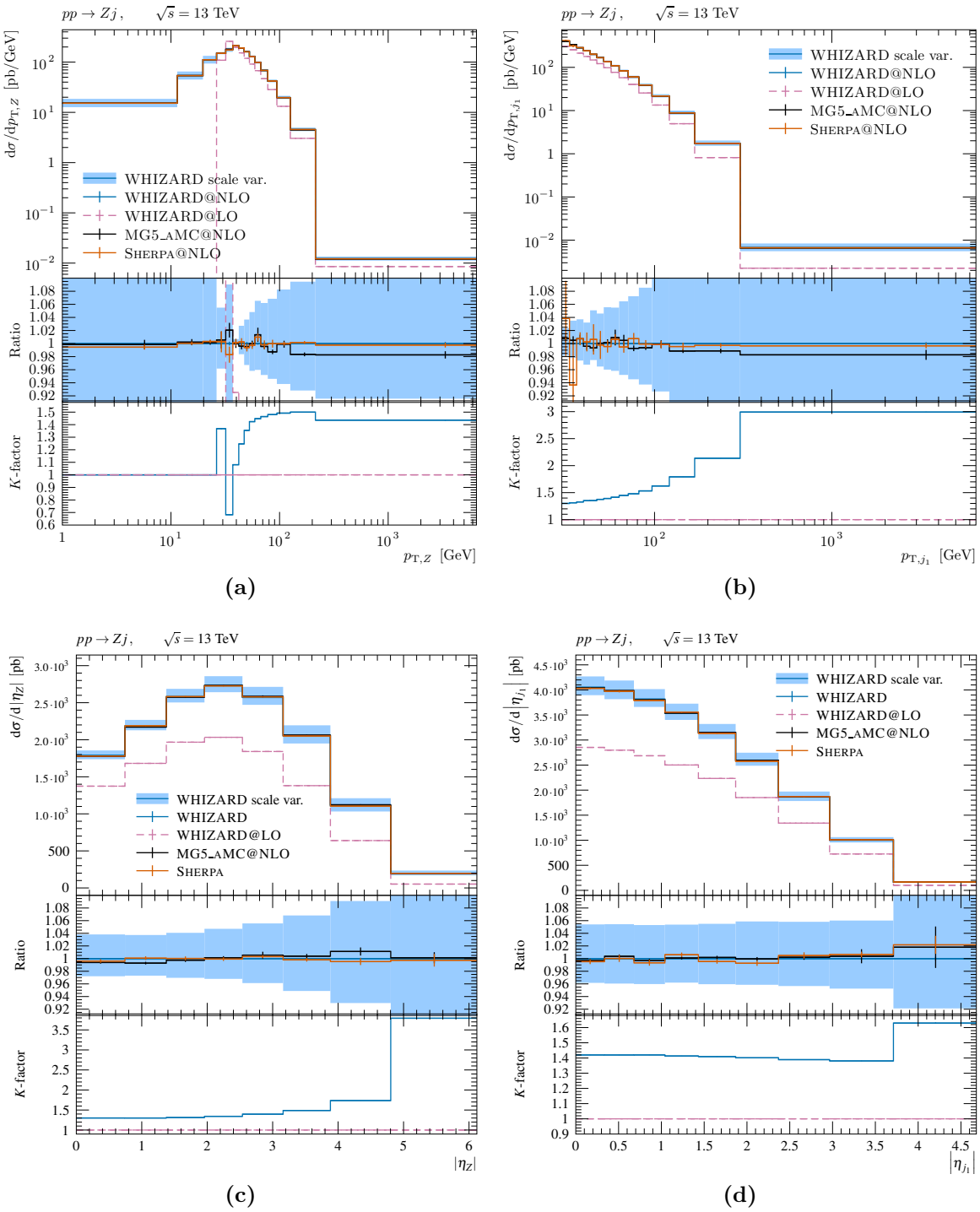


Figure 4.23. Differential distributions of transverse momenta p_T and pseudorapidities $|\eta|$ for the Z boson and hardest jet j_1 for the process $pp \to Zj$ at fNLO QCD. Statistical MC errors are shown for all three event generators, with the addition of the scale variation band for WHIZARD. The fLO distribution obtained with WHIZARD is shown as well. The ratio is shown with respect to the WHIZARD result at NLO, while the K-factor shows the NLO correction with respect to the WHIZARD LO result.

has to be a sea quark. Valence quarks – as described by the PDFs shown in fig. 4.24 as an example – have a significantly higher chance to possess a higher momentum fraction than the sea quarks and thus the Z boson is more likely to be boosted along the z-axis.

In the case where one of the quarks in the $q\bar{q} \to Z$ fusion is the result of a $g\bar{q} \leftarrow q$ (or $gq \leftarrow \bar{q}$) splitting (see fig. 2.5(a)), the initial gluon is more likely to carry a significantly different momentum fraction than the quark not stemming from the splitting, resulting in an even higher boost.

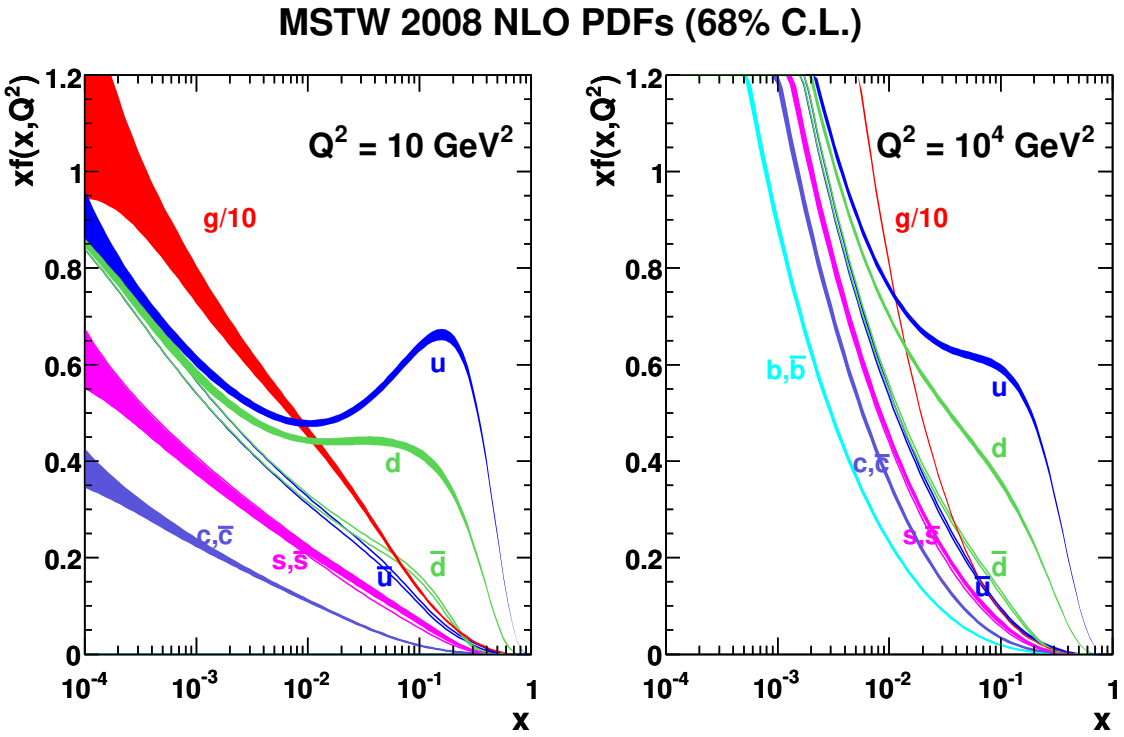


Figure 4.24. The MSTWNLO2008 PDFs dependent on the momentum fraction x of the respective parton and the factorization scale $\mu_F^2 = Q^2$. Uncertainty bands are shown for a 68% CL. The graphic is taken from [180].

In contrast to the case of $q\bar{q} \to Z$ fusion, a jet emitted from one of the incoming partons can be a soft gluon which is radiated at much higher angles to the beam axis, and thus the $|\eta_{j_1}|$ distribution has a peak close to zero. At NLO, both $|\eta|$ distributions show an almost constant K-factor across a wide region, with a strong increase for high values of $|\eta|$.

4.4. Summary and Concluding Remarks

In this chapter, we thoroughly validated our fully generalized implementation of the FKS subtraction scheme in WHIZARD for arbitrary processes at fNLO QCD. We presented examples of our methods used to assure self-consistency of the implementation within the WHIZARD framework in section 4.1. This was followed by an extensive study of the fNLO QCD capabilities of WHIZARD in comparison to two different multi-purpose MC event generators – one of which employs the CS subtraction scheme.

From our results, we conclude that the implementation to enable fully automated integration and event generation at fNLO QCD with WHIZARD is complete and validated. The main project of this thesis is thereby successfully finished, which enables many new possibilities for phenomenological studies.

We will give an example for such a study in the following chapter, where we apply the fNLO QCD capabilities of WHIZARD to multi-jet production at a lepton collider.

5. Jet Production at a Lepton Collider at Fixed-NLO QCD

With the large number of processes validated in our cross-section comparison study in section 4.2, we can confidently calculate arbitrary fNLO QCD cross sections with WHIZARD. For now, to demonstrate the possibilities of our NLO implementation, we will focus on processes with pure multi-jet final states at a future high-energy lepton collider. There, as mentioned before, precise calculations of these processes will be very important, since they not only play a part as QCD corrections to the process $e^+e^- \to f\bar{f}$, but also as a relevant background to e.g. multi-boson processes with hadronic final states. In the five-flavor scheme, this includes Higgs production with subsequent decay. Furthermore, jet multiplicities of $n_j \geq 6$ are relevant background processes to top-pair production with fully-hadronic decay. Moreover, by studying event-shape observables for multi-jet processes (as introduced in section 4.3) QCD parameters such as the strong coupling constant α_S can be probed at high precision. Any deviation from the SM predictions would give a signal for new physics in these studies.

State-of-the-art MC programs can nowadays calculate $e^+e^- \to n_j$-jet processes with jet multiplicities of up to $n_j = 5$ at fNLO QCD. For multiplicities beyond five jets, to the best of our knowledge, the only published results [206] have been calculated in the leading color approximation for up to $n_j = 7$.

Therefore, by using the excellent capability of WHIZARD to handle phase spaces of high-multiplicity final states, we extended the series of calculations of lepton-collider processes with pure colored final states consisting of $n_j = 2, \cdots, 5$ light jets (see table 4.2) to $n_j = 6$ jets. We use the same process and parameter setup as in section 4.2.1 to maintain comparability. The required OpenLoops library for the process $e^+e^- \to jjjjjj$ at NLO QCD has been generated by the OpenLoops developers on request. Our result is shown in table 5.1, together with the results for the lower jet multiplicities obtained in section 4.2.2 for convenience. As it can be seen, the result for $n_j = 6$ follows the trend of an increasing K-factor for jet multiplicities beyond $n_j = 3$. This is shown graphically in fig. 5.1 as well. The explanation for this behavior is the fact that an increase in the jet multiplicity reduces the average jet-momentum scales, leading to increasingly larger logarithms of ratios between the

CM energy and the jet-momentum scales, similar to example in eq. (4.3.15).

Using the setup from section 4.2.1 means that the result is presented for a lepton collider with a CM energy of 1 TeV. As already mentioned in the context of section 4.3, while this is higher than the realistically discussed energies of the first stage of the ILC of 250 GeV to 500 GeV [188], the technical design report [21] includes future upgrade paths up to 1 TeV. Our result thus remains relevant in the near- to mid-term future of experimental high-energy particle physics. Furthermore, it is straightforward to repeat the calculation for any reasonable CM energy.

Process	σ_{LO} [fb]	σ_{NLO} [fb]	K
$e^+e^- \to jj$	622.737(8)	639.39(5)	1.03
$e^+e^- \to jjj$	340.6(5)	317.8(5)	0.93
$e^+e^- \to jjjj$	105.0(3)	104.2(4)	0.99
$e^+e^- \to jjjjj$	22.33(5)	24.57(7)	1.10
$e^+e^- \to jjjjjj$	3.583(17)	4.46(4)	1.24

Table 5.1. Summary of fLO and fNLO QCD cross sections calculated with WHIZARD for processes with up to six light final-state jets at a lepton collider with a CM energy of 1 TeV.

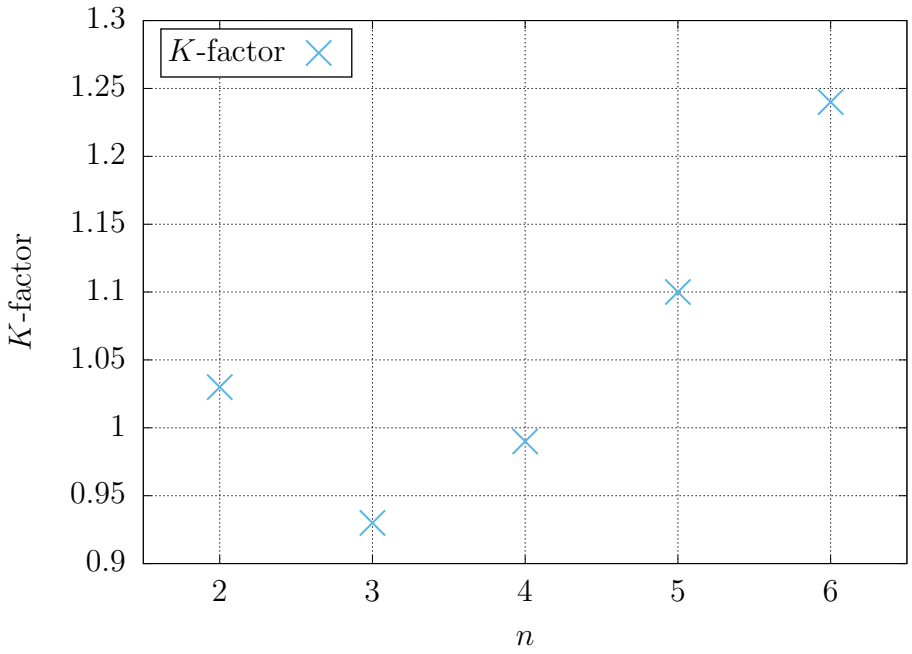


Figure 5.1. K-factor in dependence of the jet multiplicity n_j, as given in table 5.1.

Challenges and Outlook

With the first demonstration of the potential of the fNLO QCD capabilities at high final-state multiplicities of WHIZARD given in this chapter, many future studies for high-multiplicity processes at higher precision are now possible within the WHIZARD framework, providing important predictions beyond LO for future lepton colliders.

However, some challenges remain. In the context of $e^+e^- \to jjjjjj$ at fNLO QCD, future optimizations within WHIZARD regarding speed and memory consumption will be necessary. While the integration time of the Born component is fast and the time required for the real component is at the order of a few days, the computation of the 6-jet virtual-subtracted cross section has a run time at the order of several weeks to months, even with the use of the MPI-based parallelization that WHIZARD is capable of. Paired with the requirement of approximately 20 GB to 30 GB of memory per MPI instance for the OpenLoops library,[1] well-equipped computing clusters become essential for future studies of processes with high jet multiplicities.

Additionally, further optimizations to reduce a high level of redundancy in the event generation of WHIZARD are needed in order to reduce the run time for high-multiplicity processes, which usually also require a substantially higher amount of generated events for a statistically reasonable precision. So far, the cumulation of these yet unrealized optimizations unfortunately prevented us from computing differential distributions at fNLO beyond jet multiplicities of $n_j = 5$. However, while the total cross section for $n_j = 6$ is at the order of a magnitude smaller than the one for $n_j = 5$ (see table 5.1), differential distributions for up to six light jets and beyond will certainly become increasingly important in future high-precision studies.

[1] This will likely be reduced in the near future by OpenLoops thanks to a new tensor reduction method [207].

6. Summary, Conclusions, and Outlook

High-precision physics in the context of high-energy particle-collider studies plays a crucial role, with its importance ever increasing. State-of-the-art multi-purpose MC event generators allow fully-automated computations of scattering processes up to NLO accuracy, with promising computations beyond NLO becoming available for specific processes. However, a lot of work remains at NLO in perturbation theory, especially for higher multiplicities and in the context of future lepton colliders, where predictions for differential distributions at NLO remain sparse. For NLO EW and mixed corrections, fully automated implementations in MC event generators are still under active development.

With this thesis, we now provide fully automated fNLO QCD computations inside the multi-purpose MC framework WHIZARD which is well known and established in the lepton-collider community due to its multiple tools and capabilities needed to accurately simulate the various aspects of lepton collisions. In a direct comparison (i.e. without considering lepton-collider specific effects) to other multi-purpose MC event generators at fNLO QCD, WHIZARD shows excellent agreement and promising new opportunities for high-multiplicity calculations.

In part I of this thesis, we first gave a thorough review of the FKS subtraction scheme in chapter 2 for QCD in order to convey how we achieved the fNLO QCD capabilities in WHIZARD. There, we worked out the general notation used throughout this thesis. With its help we combined the various sources on FKS subtraction to give a detailed description of the construction of the universal real counterterms as introduced in eq. (2.0.2). We then laid out how they are integrated analytically over the real-emission phase space in order to cancel the divergences of the bare squared virtual amplitude $\mathcal{V}_b$ in eq. (2.0.3) as according to the KLN theorem.

Although WHIZARD has a strong focus on physics at lepton colliders, as a multi-purpose MC event generator we aim to provide a full description of SM physics at all relevant colliders. Therefore, we did not only present the FKS subtraction for FSR, but also detailed the construction of counterterms in the case of ISR due to colored initial states at a hadron collider. This also entailed additional collinear counterterms from the DGLAP evolution of the hadronic PDFs, necessary for the cancellation of

initial-state divergences in $\mathcal{V}_b$.

Following a brief general overview of WHIZARD, we gave a detailed presentation of the technical implementation of FKS subtraction in the program in chapter 3. While the basic structure for NLO computations with WHIZARD has been established in two earlier PhD theses [26, 73], it was restricted to processes only allowing for FSR of one additional gluon. We had to redesign and improve many parts of the code in order for WHIZARD to be able to compute arbitrary processes at fNLO QCD.

One notable change includes a new mapping between the flavor structure integer indices, `i_flv` and `i_hel`, and the corresponding (squared) matrix-element integer index `i` in the `state_matrix_t` type introduced in section 3.1.1 and shown in fig. 3.1.

Further into section 3.1.1, we introduced a way of optimizing the number of calls to the matrix-element generator by only requesting (squared) matrix elements for non-equivalent subprocesses. This can substantially reduce run times, especially for processes with a large number of equivalent subprocesses. So far this is only possible for matrix-element generators that provide the information which subprocesses are equivalent via the BLHA interface such as the OLP OpenLoops.

Even more importantly, we had to overhaul the way singular regions are handled in the context of FKS subtraction in WHIZARD. This then led to the correct treatment of real flavor structures that are identical except for final-state permutations, but originate from different underlying Born flavor structures, as detailed in section 3.1.3.

We further corrected symmetry factors in the real counterterms for final states with identical particles. As outlined in section 3.1.4, while these terms contain the squared Born amplitude, they require the symmetry factors of the corresponding real flavor structure. In the same manner, they require the convolution with the rescaled PDFs for the corresponding real-emission initial-state flavors and kinematics.

We successfully generalized the fNLO QCD capabilities of WHIZARD for the simulation of differential distributions as well. This included many changes in technical details regarding the generation and handling of fNLO events in WHIZARD. In section 3.1.5, we gave an overview of how to interpret fNLO events in the greater picture of the full event simulation at a real-world experiment. While they do provide physical predictions for differential distributions at FO of the perturbation series, their representation can strongly depend on the details of the technical implementation in an MC event generator due to their nature of being an intermediate step for the full event simulation. For this work, we had to determine exactly how to define FO events in the context of our implementation of FKS subtraction in WHIZARD. This was outlined in section 3.1.5 as well, together with a general description of how to treat statistical error estimates and artificial effects such as misbinning in the context of FO events.

We very briefly touched the subject of matching fNLO predictions to a parton shower.

One method that preserves fNLO accuracy for the first emission is POWHEG matching [67]. While this method has been implemented in WHIZARD for a small class of processes – i.e. lepton collisions with low-multiplicity final states with explicit (massive) quark flavors and no gluons [26, 73, 208] – the generalization of this method for arbitrary processes is an ongoing effort. We expect matched and subsequently-showered results to be presented with WHIZARD for arbitrary processes at NLO QCD in the near future, as the prerequisite of a generalized fNLO QCD framework is now fulfilled.

In part II, we first presented methods in section 4.1 to test the self-consistency of our FKS implementation and proofed that it behaves as expected on a technical level. As a next step, we conducted an fNLO QCD comparison study of total cross sections between results computed by MG5_aMC@NLO and WHIZARD, where we used [33] as a reference for our setup. We showed that WHIZARD agrees perfectly with MG5_aMC@NLO for a vast list of lepton and hadron-collider processes. Even for multi-jet processes with high multiplicities and a high complexity in FKS combinatorics, like $e^+e^- \to jjjjj$ or $pp \to jjj$, we did not observe any deviation between the results with a significance exceeding $2\sigma^{\text{sig}}_{\text{NLO}}$.[1] The deviations from MG5_aMC@NLO results for processes with one or more gluons in the final state seen in [26] have been resolved, thanks to our improved FKS implementation outlined in section 3.1.

Furthermore, both lepton and hadron-collider processes with combinations of light jets and massive gauge bosons in the final state rarely show a deviation exceeding $2\sigma^{\text{sig}}_{\text{NLO}}$, while staying below $3\sigma^{\text{sig}}_{\text{NLO}}$. The only noteworthy deviation we encountered was for the process $pp \to W^+W^-ZZ(4f)$ with $3.56\sigma^{\text{sig}}_{\text{NLO}}$. We discussed several possible explanations for this deviation, including possible differences in the phase-space treatments between the two programs, a possible error in the OpenLoops library, or simply an overlooked error in the process setup for one of the programs. While this deviation warrants further investigations, the significance is still rather low and it is a single outlier in a vast list of processes without any significant deviations between MG5_aMC@NLO and WHIZARD.

To fully compare the list in [33], processes with final-state Higgs bosons or single top quarks, as well as processes requiring photon isolation or b-jet tagging, will have to be computed with WHIZARD. While we do expect the former two classes of processes to be free of surprises, the latter two require additional validation of the photon isolation and b-jet tagging that have only recently been implemented in WHIZARD. However, these methods are not exclusive to NLO calculations.

In section 4.3, we conducted a thorough comparison study between MG5_aMC@NLO, Sherpa, and WHIZARD for differential distributions of various observables at fNLO QCD in multi-jet production at a future lepton collider with a CM energy of 1 TeV.

[1] See eq. (4.2.8) for the definition of $\sigma^{\text{sig}}_{\text{NLO}}$

For a jet multiplicity of up to $n_j = 5$ at Born level, we studied the agreement for different observables of jet properties – i.e. their transverse momenta, energies, absolute pseudorapidities, and invariant masses – and for multiple different types of event-shape observables – i.e. thrust and thrust major, oblateness, hemisphere broadenings and masses, and Parisi variables.

We found an excellent agreement between all three MC event generators for jet multiplicities of up to $n_j = 4$. For $n_j = 5$, MG5_aMC@NLO shows multiple bin-wise statistical deviations of $5\sigma^{\text{sig}}_{\text{NLO}}$ and above in relation to Sherpa and WHIZARD for several observables. Because of the rather long run time of roughly a month for 5-jet production at fNLO with MG5_aMC@NLO, we were not able to determine the origin of this discrepancy in time for this thesis. We made the point, however, that the deviations would likely become insignificant when scale variations of both MG5_aMC@NLO and WHIZARD would be considered. Additionally, the fact that Sherpa – which employs the CS subtraction scheme – agrees well with WHIZARD even for $n_j = 5$ gives us further reason to remain confident in the fNLO QCD capabilities of WHIZARD.

In order to present the capabilities of WHIZARD to produce differential distributions for hadron-collider processes at fNLO QCD, we showed distributions for transverse momenta and absolute pseudorapidities for the process $pp \to Zj$ in section 4.3.5. We saw an excellent agreement between MG5_aMC@NLO, Sherpa, and WHIZARD.

From our comparison study of multi-purpose MC event generators, we conclude that automated fNLO QCD computations with WHIZARD are correct and validated, and are now ready for phenomenological studies. Our implementation allows to combine the fNLO QCD capabilities with the various advantages of the WHIZARD framework, such as the support of lepton-collider beam polarization and spectra, QED ISR, as well as its excellent phase-space setup and speed in integration and event generation.

Finally, with a high confidence in our FKS implementation due to the results stated above, we applied the outstanding capability of WHIZARD to describe phase spaces for high-multiplicity final states to 6-jet production at fNLO QCD at a future 1 TeV lepton collider. As far as we know, we computed the first result for the total cross section in such a setting. The result of $\sigma_{\text{NLO}} = 4.46(4)$ fb fits well into the expected range of values we would infer from computations of increasing jet multiplicities in table 5.1. The relatively high K-factor of 1.24 implies that FO perturbation theory might not be sufficient for increasing jet multiplicities as large logarithms start to enhance the fNLO correction. These logarithms could be resummed by matching the fNLO predictions to a parton shower – e.g. achieved by the `POWHEG` method – in order to give more accurate predictions for experiments.

This thesis has focussed on fNLO QCD corrections. In the context of high-precision studies, especially at a future lepton collider, including additional effects from higher-

order EW and mixed corrections in theoretical predictions will become increasingly important. We implemented a first iteration of fNLO EW corrections within our FKS framework. However, we refrained from presenting it in this thesis, as it is still work in progress and initial validation efforts have shown promising results only very recently. We are hopeful that first results at fNLO EW will be published with WHIZARD in the near future. The implementation of mixed corrections should be possible by doubling the component structure of WHIZARD presented in section 3.1.1 such that both QCD and EW corrections can be combined accordingly.

As stated in [26, 73], NNLO computations should in principle be possible by this component structure as well. While the dipole-based antenna subtraction scheme has gotten a lot of attention in the recent years, an FKS-inspired NNLO approach as realized in the STRIPPER subtraction scheme [209, 210] would be preferred for the implementation in WHIZARD as proposed in [26, 73]. A very recent example of an application of this scheme to diphoton production with an additional jet at the LHC is given in [211]. In order to accommodate the additional terms for the double-real, double-virtual, and real-virtual contributions, corresponding components could simply be allocated in addition to the ones needed at NLO [26, 73]. Real-virtual and tree-level squared amplitudes can be provided already by some matrix-element generators like OpenLoops, while the two-loop computations remain to be fully automated. It is very probable that this task will be resolved by the time NNLO computations will be implemented in WHIZARD, as it has become a very active field of research in the recent years.

A. Real-radiation Phase-space Construction

In chapter 2, for an arbitrary hard $2 \to n$ scattering process, we introduced the FKS subtraction scheme with a general description for the parametrization of the real-emission phase space and its associated set of kinematics Φ_{n+1}. There, in eq. (2.1.15) we introduced the injective mapping between Φ_{n+1}, and the barred kinematics $\bar{\Phi}_n^{(\alpha)}$ of the underlying Born configuration of a singular region α and its associated set of FKS variables $\Phi_{\text{rad}}^{(\alpha)} = \{\xi, y, \phi\}$ that parametrize the kinematics of the emitted FKS parton.

While the details of the phase-space construction for the real radiation did not matter for the correct derivation of the FKS subtraction terms, the implementation of the FKS subtraction scheme in a multi-purpose MC event generator of course requires an explicit generation of real-emission kinematics. The construction of the $\bar{\Phi}_n^{(\alpha)}$ and $\Phi_{\text{rad}}^{(\alpha)}$ kinematics from the Φ_{n+1} kinematics has been derived in [67] and are reviewed in detail in [212]. The inverse construction, i.e. the construction of the full set Φ_{n+1} from a given set of barred kinematics $\bar{\Phi}_n^{(\alpha)}$ and a given set of FKS variables $\Phi_{\text{rad}}^{(\alpha)}$, has been computed in [67] as well.

For the FKS implementation in WHIZARD, the inverse construction is the preferred and most convenient one, as then the barred kinematics $\bar{\Phi}_n^{(\alpha)}$ simply equate to that of the Born phase space and are readily available from the existing routines to generate the n-body phase space at LO. The details of the inverse construction in the context of WHIZARD is documented in [26, 212].

For completeness, in the following we review the construction of the real-radiation kinematics as described in [67] for FSR in appendix A.1, as well as for ISR in appendix A.2. In appendix A.3 the jacobians $\mathcal{J}^{(\alpha)}$ for the real-radiation phase-space transformation are given. Throughout this appendix we use the notation established in section 2.1.

A.1. Final-state Real-radiation Phase Space

First, we will focus on FSR from a massless emitter and defer the treatment of massive emitters to appendix A.1.1. We will detail the inverse construction of the $(n+1)$-body kinematics as given in [67]. In the case of FSR, the initial-state four-momenta are

unaffected by the real emission, thus

$$k_\oplus = \bar{k}_\oplus \,, \quad k_\ominus = \bar{k}_\ominus \,, \tag{A.1.1}$$

which consequently leads to the same CM frame for the sets of final-state four-momenta

$$\bar{\mathcal{F}} \equiv \left\{\bar{k}_3, \cdots, \bar{k}_{n+2}\right\} \,, \quad \mathcal{F} \equiv \left\{k_3, \cdots, k_{n+3}\right\} \,, \tag{A.1.2}$$

associated to the n-body and $(n+1)$-body phase space, respectively. Therefore, for the FSR phase-space construction, we work in this CM frame. Obviously, momentum conservation has to be satisfied such that

$$q \equiv \bar{k}_\oplus + \bar{k}_\ominus = \sum_{i=3}^{n+2} \bar{k}_i = k_\oplus + k_\ominus = \sum_{i=3}^{n+3} k_i \,, \tag{A.1.3}$$

where we defined q, which is given in the CM frame as[1]

$$\mathbf{q} = 0 \,, \quad q^2 = \left(q^0\right)^2 \equiv s \,. \tag{A.1.4}$$

We also assume, without loss of generality, that we have a single singular region that is associated with the FKS pair $(n+2, n+3)$ (see eq. (2.1.11)), i.e. the emitted parton $\mathcal{I}_{n+3}$ becoming soft and/or collinear to the emitting parton $\mathcal{I}_{n+2}$. We thus suppress the index α. The construction of $\mathcal{F}$ is then based on the separation of the final-state real-emission kinematics into two systems such that

$$\mathcal{F} = \mathcal{C} \cup \mathcal{E} \,, \quad \mathcal{C} \equiv \left\{k_3, \cdots, k_{n+1}\right\} \,, \quad \mathcal{E} \equiv \left\{k_{n+2}, k_{n+3}\right\} \,, \tag{A.1.5}$$

where $\mathcal{E}$ are the kinematics of the emitter-emitted system and $\mathcal{C}$ that of the recoiling system. We also introduce the sum of four-momenta of $\mathcal{E}$ as

$$k \equiv k_{n+2} + k_{n+3} \,. \tag{A.1.6}$$

The mapping between $\bar{\mathcal{F}}$ and $\mathcal{F}$ is constructed in such a way that the invariant mass M_{rec} of the recoiling system $\mathcal{C}$ is preserved, meaning that

$$\bar{k}_{\text{rec}}^2 = \left(\sum_{i=3}^{n+1} \bar{k}_i\right)^2 \overset{!}{=} \left(\sum_{i=3}^{n+1} k_i\right)^2 = k_{\text{rec}}^2 = M_{\text{rec}}^2 \,, \tag{A.1.7}$$

[1] This is the default definition. In more general formulations of FKS, q^0 is defined as the energy of a subset of all particles in the final state. For example, the resonance-aware FKS subtraction scheme [69] defines q^0 as the resonance energy of the associated resonance history.

with the definition

$$k_{\text{rec}} \equiv q - k\,, \quad \mathbf{k}_{\text{rec}} = -\mathbf{k}\,. \tag{A.1.8}$$

Equation (A.1.7) implies that the four-momenta of $\mathcal{C}$ and their respective barred counterparts of the set

$$\bar{\mathcal{C}} \equiv \left\{\bar{k}_3, \cdots, \bar{k}_{n+1}\right\}\,, \tag{A.1.9}$$

are connected via a boost Λ along the $\mathbf{k}_{\text{rec}}$ direction such that

$$(q - \Lambda k_{\text{rec}})^2 = 0\,, \tag{A.1.10}$$

i.e. the difference $q - \Lambda k_{\text{rec}}$ is light-like. We then have

$$\bar{k}_i = \Lambda k_i\,, \quad i = 3, \ldots, n+1\,, \quad \bar{k}_{n+2} = q - \Lambda k_{\text{rec}}\,, \tag{A.1.11}$$

where momentum conservation for the sum of all barred momenta as in eq. (A.1.3) is obviously satisfied. From eqs. (A.1.3), (A.1.7) and (A.1.11) it follows that

$$M^2_{\text{rec}} = \left(q - \bar{k}_{n+2}\right)^2\,. \tag{A.1.12}$$

With these basic considerations, the inverse construction of $\mathcal{C}$ and $\mathcal{E}$ from the barred four-momenta $\bar{\mathcal{F}}$ then is rather straightforward and depends on the set of FKS variables (see eqs. (2.1.19) and (2.1.21) to (2.1.23))

$$\Phi_{\text{rad}} = \{\xi, y, \phi\} \tag{A.1.13}$$

that parametrize the kinematics of the emitted parton in the CM frame. Within this parametrization, the energy of the emitted massless parton $\mathcal{I}_{n+3}$ is given as

$$k^0_{n+3} = |\mathbf{k}_{n+3}| = \frac{q^0}{2}\xi\,, \tag{A.1.14}$$

where we also stated the trivially given magnitude $|\mathbf{k}_{n+3}|$ of the three-momentum of $\mathcal{I}_{n+3}$. Whats left for the construction of k_{n+3} is to figure out its polar and azimuthal angles. In order to do so, we first have to compute the magnitudes $|\mathbf{k}_{n+2}|$ and $|\mathbf{k}|$. For a massless emitter we have

$$y \equiv \cos\theta = \frac{\mathbf{k}_{n+2}\cdot\mathbf{k}_{n+3}}{|\mathbf{k}_{n+2}||\mathbf{k}_{n+3}|}\,, \tag{A.1.15}$$

which gives

$$|\mathbf{k}|^2 = |\mathbf{k}_{n+2}|^2 + |\mathbf{k}_{n+3}|^2 + 2|\mathbf{k}_{n+2}||\mathbf{k}_{n+3}|y\,. \tag{A.1.16}$$

Energy conservation gives us

$$q^0 = k^0_{n+2} + k^0_{n+3} + k^0_{\text{rec}} = |\mathbf{k}_{n+2}| + |\mathbf{k}_{n+3}| + \sqrt{|\mathbf{k}|^2 + M^2_{\text{rec}}}\,, \tag{A.1.17}$$

where we used the fact that because of momentum conservation we have $|\mathbf{k}_{\text{rec}}| = |\mathbf{k}|$, and that both emitter and emitted parton are massless, giving

$$k^0_{n+2/3} = \left|\mathbf{k}_{n+2/3}\right|. \tag{A.1.18}$$

Here, we introduced the notation

$$\mathcal{I}_{n+2/3} \tag{A.1.19}$$

as a shortcut to simultaneously describe emitter and emitted parton. We note that the order of the integer before and after the slash matters. Squaring eq. (A.1.17) and inserting eq. (A.1.16) leads to

$$|\mathbf{k}_{n+2}| = k^0_{n+2} = \frac{q^2 - M^2_{\text{rec}} - 2q^0 k^0_{n+3}}{2\left(q^0 - k^0_{n+3}\,(1-y)\right)}\,. \tag{A.1.20}$$

We thus have determined both emitter and emitted four-momenta $k_{n+2/3}$, except for their polar and azimuthal angles. However, we can now position their three-momenta parallel to the barred emitter momentum $\bar{\mathbf{k}}_{n+2}$ such that

$$\mathbf{k}''_{n+2/3} \equiv \left|\mathbf{k}_{n+2/3}\right| \frac{\bar{\mathbf{k}}_{n+2}}{\left|\bar{\mathbf{k}}_{n+2}\right|}\,, \tag{A.1.21}$$

and apply a rotation $R(\psi_{n+2/3}, \mathbf{d})$ around the unit vector $\mathbf{d} \perp \bar{\mathbf{k}}_{n+2}$ to both three-momenta $\mathbf{k}''_{n+2/3}$ such that

$$\mathbf{k}'_{n+2/3} = R(\psi_{n+2/3}, \mathbf{d})\mathbf{k}''_{n+2/3}\,, \quad \cos\psi_{n+2/3} \equiv \frac{\mathbf{k}_{n+2/3}\cdot\mathbf{k}}{\left|\mathbf{k}_{n+2/3}\right||\mathbf{k}|}\,, \tag{A.1.22}$$

with

$$\mathbf{d} \equiv \frac{1}{\sqrt{\left(\bar{k}^1_{n+2}\right)^2 + \left(\bar{k}^2_{n+2}\right)^2}} \left(\bar{k}^2_{n+2}, -\bar{k}^1_{n+2}, 0\right)^T. \tag{A.1.23}$$

Without knowing the direction of $\mathbf{k}$, we can use momentum conservation

$$\left(\mathbf{k} - \mathbf{k}_{n+2/3}\right)^2 = \mathbf{k}^2_{n+3/2} \tag{A.1.24}$$

to obtain

$$\cos\psi_{n+2/3} = \frac{\left|\mathbf{k}_{n+2/3}\right|^2 + |\mathbf{k}|^2 - \left|\mathbf{k}_{n+3/2}\right|^2}{2\left|\mathbf{k}_{n+2/3}\right||\mathbf{k}|}\,. \tag{A.1.25}$$

The only missing step to fully construct the emitter-emitted system $\mathcal{E}$ then is to rotate the primed three-momenta by the azimuthal angle $\phi \in \Phi_{\text{rad}}$, which yields the three-momenta of $\mathcal{E}$ as

$$\mathbf{k}_{n+2/3} = R\left(\phi, \frac{\bar{\mathbf{k}}_{n+2}}{|\bar{\mathbf{k}}_{n+2}|}\right) \mathbf{k}'_{n+2/3}\,. \tag{A.1.26}$$

In order to restore total momentum conservation, we still have to boost the recoil system $\mathcal{C}$ via the inverse Lorentz transformation $\Lambda^{-1}(\beta)$ (see eq. (A.1.11)) along the $-\mathbf{k}_{\text{rec}}$ direction, dependent on the boost velocity β. The boost velocity can be easily derived by considering eq. (A.1.10), which directly leads to the condition

$$q^0 - \gamma k^0_{\text{rec}} - \gamma\beta|\mathbf{k}_{\text{rec}}| \stackrel{!}{=} \gamma\beta k^0_{\text{rec}} + \gamma|\mathbf{k}_{\text{rec}}|\,. \tag{A.1.27}$$

Solving for q^0 and using the square $(q^0)^2 = q^2$ gives

$$q^2 = \frac{1+\beta}{1-\beta}\left(k^0_{\text{rec}} + |\mathbf{k}_{\text{rec}}|\right)^2\,, \tag{A.1.28}$$

which in turn, by solving for β, leads to

$$\beta = \frac{q^2 - \left(k^0_{\text{rec}} + |\mathbf{k}_{\text{rec}}|\right)^2}{q^2 + \left(k^0_{\text{rec}} + |\mathbf{k}_{\text{rec}}|\right)^2}\,. \tag{A.1.29}$$

We note that because of the fact that $q^0 = k^0 + k^0_{\text{rec}}$, and that the invariant mass of $\mathcal{E}$ is $k^2 \geq 0$, giving $k^0 \geq |\mathbf{k}| = |\mathbf{k}_{\text{rec}}|$, the boost velocity β is positive and smaller than one. This ensures that the boost $\Lambda(\beta)$ always exists. The recoil system $\mathcal{C}$ is thus fully constructed by

$$k_i = \Lambda^{-1}(\beta)\bar{k}_i\,, \quad i = 3, \cdots, n+1\,. \tag{A.1.30}$$

The inverse mapping of the real kinematics from the barred set of momenta $\bar{\Phi}_n$ and the FKS variables Φ_{rad} thus is always unique and exists, provided that

$$\xi < \frac{q^2 - M^2_{\text{rec}}}{q^2}\,, \tag{A.1.31}$$

as $|\mathbf{k}_{n+2}| > 0$. With eq. (A.1.12), and $\bar{\mathbf{k}}^2_{n+2} = 0$ for a massless emitter, eq. (A.1.31) yields the maximum radiation energy

$$\xi_{\text{max}} = \frac{2|\bar{\mathbf{k}}_{n+2}|}{q^0}\,. \tag{A.1.32}$$

A.1.1. Construction for Massive Emitters

The construction of the $\Phi_{n+1} = \{k_\oplus, k_\ominus, k_3, \cdots, k_{n+3}\}$ kinematics in the case of a massive final-state emitter $\mathcal{I}_{n+2}$ with mass m is derived in [68], with a detailed review in [212]. It is in many parts analogous to the construction in the massless case, detailed above in appendix A.1. However, due to the massive emitter, we have a different construction for the three-vector magnitude $|\mathbf{k}_{n+2}|$ and energy k^0_{n+2} of the four-momentum of the emitter in the $(n+1)$-kinematics, as well as for the magnitude $|\mathbf{k}|$ of the sum of momenta of the emitter-emitted system $\mathcal{E}$. These quantities determine the rotation given in eq. (A.1.22), affecting the construction of $\mathcal{E}$. We will give a short review of their derivation as given in [68] for a massive final-state emitter in the following.

From eqs. (A.1.3), (A.1.4) and (A.1.7) and

$$\bar{k}^2_{n+2} = \left(q - \bar{k}_{\text{rec}}\right)^2 = m^2\,, \quad \bar{k}^2_{\text{rec}} = \left(q - \bar{k}_{n+2}\right)^2 = M^2_{\text{rec}}\,, \tag{A.1.33}$$

we can write

$$\bar{k}^0_{n+2} = \frac{q^2 + m^2 - M^2_{\text{rec}}}{2q^0}\,, \quad \bar{k}^0_{\text{rec}} = \frac{q^2 - m^2 + M^2_{\text{rec}}}{2q^0}\,. \tag{A.1.34}$$

While for the massless emitted parton $\mathcal{I}_{n+3}$ the energy and magnitude is simply given by eq. (A.1.14), the construction of these quantities for the four-momentum k_{n+2} of the massive emitter relies on the established fact in [68] that k^0_{n+2}, k^0_{n+3}, and k^0_{rec} exist in a convex Dalitz domain. This allows to introduce the parametrization

$$k^0_{n+2} = \bar{k}^0_{n+2} - z k^0_{n+3}\,. \tag{A.1.35}$$

There then exists a maximum value of $k^0_{n+3} = |\mathbf{k}_{n+3}|$ for any value of z. This maximum is such that it can be associated with a point on the boundary of the Dalitz region that is determined by at least one sign combination of the condition

$$|\mathbf{k}_{n+3}| \pm |\mathbf{k}_{n+2}| \pm |\mathbf{k}_{\text{rec}}| = 0\,. \tag{A.1.36}$$

By squaring two times such that

$$\left(|\mathbf{k}_{n+3}| \pm |\mathbf{k}_{n+2}|\right)^2 = |\mathbf{k}_{n+3}|^2 + |\mathbf{k}_{n+2}|^2 \pm 2|\mathbf{k}_{n+3}||\mathbf{k}_{n+2}| = |\mathbf{k}_{\text{rec}}|^2\,, \tag{A.1.37}$$

and subsequently

$$\left(|\mathbf{k}_{n+3}|^2 + |\mathbf{k}_{n+2}|^2 - |\mathbf{k}_{\text{rec}}|^2\right)^2 = 4|\mathbf{k}_{n+3}|^2|\mathbf{k}_{n+2}|^2\,, \tag{A.1.38}$$

we can then make use of eqs. (A.1.34) and (A.1.35), as well as

$$|\mathbf{k}_{n+2}|^2 = \left(k^0_{n+2}\right)^2 - m^2\,, \quad |\mathbf{k}_{\rm rec}|^2 = \left(q^0 - k^0_{n+2} - k^0_{n+3}\right)^2 - M^2_{\rm rec}\,, \tag{A.1.39}$$

to write

$$4|\mathbf{k}_{n+3}|^2 \left(2|\mathbf{k}_{n+3}|q^0 z\,(1-z) + q^2 z^2 - 2q^0 \bar{k}^0_{\rm rec} z + M^2_{\rm rec}\right) = 0\,. \tag{A.1.40}$$

Solving this quadratic equation for $|\mathbf{k}_{n+3}| = k^0_{n+3}$ gives $k^0_{n+3} = 0$ and its maximum value such that

$$0 \leq k^0_{n+3} \leq \frac{2q^0 \bar{k}^0_{\rm rec} z - q^2 z^2 - M^2_{\rm rec}}{2q^0 z\,(1-z)}\,. \tag{A.1.41}$$

The minimum and maximum values of z are those for which the rightmost expression in eq. (A.1.41) vanishes, which turn out to be

$$z_{1/2} = \frac{\bar{k}^0_{\rm rec} \pm \sqrt{\left(\bar{k}^0_{\rm rec}\right)^2 - M^2_{\rm rec}}}{q^0}\,, \tag{A.1.42}$$

which we can use to parametrize $z_1 < z < z_2$ as

$$z = z_2 - \frac{1+y}{2}\,(z_2 - z_1)\,. \tag{A.1.43}$$

Thus, for the case of a massive emitter, we exchange y with z in $\Phi_{\rm rad}$. Together with eqs. (A.1.35) and (A.1.39), and $|\mathbf{k}| = |\mathbf{k}_{\rm rec}|$ (see eqs. (A.1.6) and (A.1.8)), as well as

$$|\mathbf{k}_{n+3}| = k^0_{n+3} = \frac{q^0}{2}\xi\,, \tag{A.1.44}$$

we have all the quantities needed to perform the construction of $\mathcal{E}$ for a massive emitter. It is then simply a matter of inserting these quantities in eqs. (A.1.21) to (A.1.26). Furthermore, by inserting the maximum of k^0_{n+3} from eq. (A.1.41) into eq. (A.1.44), we derive the maximum value of ξ as

$$\xi_{\rm max} = \frac{2}{q^0}\frac{2q^0 \bar{k}^0_{\rm rec} z - q^2 z^2 - M^2_{\rm rec}}{2q^0 z\,(1-z)}\,. \tag{A.1.45}$$

For the recoiling momenta $\mathcal{C}$, we then perform an inverse Lorentz boost $\Lambda^{-1}(\beta)$ along the direction of $-\mathbf{k}_{\rm rec} = \mathbf{k}$ as in eq. (A.1.30), where the boost velocity is given as [26]

$$\beta = \frac{1-\alpha^2}{1+\alpha^2}\,, \quad \alpha \equiv \frac{k^0_{\rm rec} + |\mathbf{k}_{\rm rec}|}{k^0_{\rm rec} - |\mathbf{k}_{\rm rec}|}\,. \tag{A.1.46}$$

A.2. Initial-state Real-radiation Phase Space

Initial-state QCD radiation in general has to be treated in the context of PDFs that determine the partonic momenta of the colliding beams (see section 2.2.2). This complicates the construction of the kinematics for ISR. However, in general, we can assume the beam partons to be sufficiently fast enough to not have their momentum direction altered by radiation. Thus, while ISR affects their energy, their momentum direction can be assumed to be left parallel to the beam direction. As before, we follow the construction of the kinematics as given in [68]. We assume, without loss of generality, that we have a single singular region, where the emitted parton is $\mathcal{I}_{n+3}$, i.e. its four-momentum is positioned last in the set

$$\Phi_{n+1} = \{k_\oplus, k_\ominus, k_3, \cdots, k_{n+3}\} \tag{A.2.1}$$

of the $(n+1)$-body kinematics. We thus suppress the index α. In terms of the FKS variables $\Phi_{\text{rad}} = \{\xi, y, \phi\}$ (see eqs. (2.1.19) and (2.1.21) to (2.1.23)), the four-momentum k_{n+3} of $\mathcal{I}_{n+3}$, emitted from an initial-state parton $\mathcal{I}_\oplus$, is then given by

$$k_{n+3} = \frac{\sqrt{s}}{2}\xi\left(1, \sqrt{1-y^2}\sin\phi, \sqrt{1-y^2}\cos\phi, y\right), \tag{A.2.2}$$

where $y = \cos\theta_{\oplus j}$ as in eq. (2.1.22). Equation (A.2.2) and the following calculations are set in the CM frame of the partonic initial-state system of the $(n+1)$-body process, where

$$q \equiv \left(\sqrt{s}, \mathbf{0}\right) = k_\oplus + k_\ominus = \sum_{i=3}^{n+3} k_i\,, \quad k_\oplus = x_\oplus K_\oplus\,, \tag{A.2.3}$$

with the initial-state momenta written in terms of the beam momenta $K_\oplus$ and the partonic momentum fractions $x_\oplus$. We introduce the total momentum $\bar{k}_{\text{tot}}$ of the final state of the underlying Born process, and choose the mapping between the real kinematics and the barred momenta, as well as the FKS variables, in such a way that

$$\bar{k}_{\text{tot}}^2 = \left(\sum_{i=3}^{n+2}\bar{k}_i\right)^2 \overset{!}{=} \left(\sum_{i=3}^{n+2} k_i\right)^2 = k_{\text{tot}}^2\,, \tag{A.2.4}$$

$$\bar{k}_{\text{tot}} \equiv \sum_{i=3}^{n+2}\bar{k}_i\,, \quad k_{\text{tot}} \equiv \sum_{i=3}^{n+2} k_i = k_\oplus + k_\ominus - k_{n+3}\,, \tag{A.2.5}$$

meaning that the invariant mass of $\bar{k}_{\text{tot}}$ and its equivalent k_{tot} in the $(n+1)$-body final state is preserved under Lorentz transformations that connect these two four-momenta. With the additional condition that the rapidity of $\bar{k}_{\text{tot}}$ and k_{tot} has to be the same, we construct a series of Lorentz boosts, where the first is a boost Λ_L longitudinal with

respect to the incoming beams such that the rapidity of both $\Lambda_L \bar{k}_{\text{tot}}$ and $\Lambda_L k_{\text{tot}}$ vanishes. The boost angle of Λ_L is given by minus the rapidity of $\bar{k}_{\text{tot}}$, with the components of the boost velocity β_L given as [26]

$$\beta_L^i = \frac{\bar{k}_{\text{tot}}^i}{\bar{k}_{\text{tot}}^0} = \frac{\bar{k}_\oplus^i + \bar{k}_\ominus^i}{\bar{k}_\oplus^0 + \bar{k}_\ominus^0} \,. \tag{A.2.6}$$

A second transverse boost Λ_T assures that $\Lambda_T \Lambda_L k_{\text{tot}}$ has no transverse component and is constructed such that

$$\Lambda_T \Lambda_L k_{\text{tot}} = \Lambda_L \bar{k}_{\text{tot}} = \sum_{i=3}^{n+2} \Lambda_L \bar{k}_i \,, \tag{A.2.7}$$

with the boost velocity β_T given as [26]

$$\beta_T = \left(1 + \frac{s\,(1-\xi)}{k_{T,n+3}^2}\right)^{-\frac{1}{2}} \,, \tag{A.2.8}$$

where $k_{T,n+3}$ is the transverse momentum of the emitted parton with respect to the beam direction (z-axis). From eq. (A.2.7) we then immediately obtain the inverse construction of the rest of the $(n+1)$-body final-state four-momenta from the barred momenta as

$$k_i = \Lambda_L^{-1} \Lambda_T^{-1} \Lambda_L \bar{k}_i \,, \quad i = 3, \cdots, n+2 \,, \tag{A.2.9}$$

with the four-momentum of the emitted parton already given in eq. (A.2.2). For the construction of the $(n+1)$-body initial-state momenta we then define the barred parton-momentum fractions $\bar{x}_\oplus$ such that

$$\bar{x}_\oplus K_\oplus + \bar{x}_\ominus K_\ominus = \bar{k}_{\text{tot}} \tag{A.2.10}$$

where we can use $\bar{x}_\oplus$ to compute $x_\oplus$ and thus $k_\oplus = x_\oplus K_\oplus$ since [67]

$$x_\oplus = \frac{\bar{x}_\oplus}{\sqrt{1-x}} \sqrt{\frac{2 - \xi\,(1 \mp y)}{2 - \xi\,(1 \pm y)}} \,. \tag{A.2.11}$$

This relation follows directly from the fact that the invariant mass and rapidity is kept equal for k_{tot} and $\bar{k}_{\text{tot}}$, with its derivation being detailed in [212]. By considering the constraint that $x_\oplus \leq 1$, we can work out an upper bound ξ_{max} on ξ in the case of ISR by first squaring eq. (A.2.11) and subsequently adding $0 = y - y$ to the numerator and

denominator of the second fraction such that

$$x_{\oplus}^2 = \frac{\bar{x}_{\oplus}^2}{1-\xi}\frac{(1\pm y)+(1-\xi)(1\mp y)}{(1\mp y)+(1-\xi)(1\pm y)} \leq 1\,. \tag{A.2.12}$$

Transferring this into a quadratic equation in $(1-\xi)$ leads to

$$(1-\xi)^2 + (1-\xi)\frac{(1\mp y)\left(1-\bar{x}_{\oplus}^2\right)}{1\pm y} - \bar{x}_{\oplus}^2 \geq 0\,. \tag{A.2.13}$$

With the obvious condition $1-\xi \geq 0$ we get the two solutions

$$1-\xi = \frac{2(1\pm y)\,\bar{x}_{\oplus}^2}{(1\mp y)\left(1-\bar{x}_{\oplus}^2\right) + \sqrt{\left(1+\bar{x}_{\oplus}^2\right)^2(1\mp y)^2 + 16y\bar{x}_{\oplus}^2}}\,. \tag{A.2.14}$$

The upper bound $\xi_{\max}$ is then given by the larger of the two solutions, corresponding to the lower value of ξ such that

$$\xi_{\max} = 1 - \max\left\{\frac{2(1+y)\,\bar{x}_{\oplus}^2}{(1-y)(1-\bar{x}_{\oplus}^2) + \sqrt{(1+\bar{x}_{\oplus}^2)^2(1-y)^2 + 16y\bar{x}_{\oplus}^2}}, \frac{2(1-y)\,\bar{x}_{\ominus}^2}{(1+y)(1-\bar{x}_{\ominus}^2) + \sqrt{(1+\bar{x}_{\ominus}^2)^2(1+y)^2 + 16y\bar{x}_{\ominus}^2}}\right\}. \tag{A.2.15}$$

A.3. The Jacobian for the Real-radiation Phase-space Transformation

The Jacobian $\mathcal{J}^{(\alpha)}$, arising from the variable transformation when expressing the radiation phase space in terms of $\Phi_{\text{rad}}^{(\alpha)}$ and needed to compute the radiation phase-space measure $d\Phi_{\text{rad}}^{(\alpha)}$ (see eq. (2.1.18)) in a singular phase-space region α, is derived in [67] for ISR and FSR with massless emitters. For massive final-state emitters, it is derived in [68]. In [212], the derivation of $\mathcal{J}^{(\alpha)}$ is reviewed in additional detail. In the following, we will simply state their final form, together with the associated phase-space measure $d\Phi_{\text{rad}}^{(\alpha)}$, for the different cases. We assume a general FKS pair $(i,j) \in \mathcal{P}_{\text{FKS}}$ (see eq. (2.1.11)).

- FSR, $m_i = 0$, $i > n_I \mid (i,j) \in \mathcal{P}_{\text{FKS}}$:

  In the case of FSR with a massless emitter $\mathcal{I}_i$, we have [67]

$$d\Phi_{\text{rad}}^{(\alpha)} = \mathcal{J}^{(\alpha)} d\xi_j dy_{ij} d\phi_j \,, \quad \mathcal{J}^{(\alpha)} = \frac{q^2 \xi_j}{(4\pi)^3} \frac{k_i^2}{\bar{k}_i^0} \left(k_i^0 - \frac{k^2}{2q^0} \right)^{-1} \,, \qquad k = 2k_i^0 k_j^0 \left(1 - y_{ij}\right) \,. \tag{A.3.1}$$

  In the soft limit, we have $\xi_j \to 0$, $k_j^0 \to 0$, and thus $k \to 0$. Consequently, the soft limit of eq. (A.1.20) is

$$\lim_{\xi_j \to 0} k_i^0 = \frac{q^2 - M_{\text{rec}}}{2q^0} = \bar{k}_i^0 \,, \tag{A.3.2}$$

  with the last equality being because of eq. (A.1.12). We thus have

$$\lim_{\xi_j \to 0} \mathcal{J}^{(\alpha)} = \frac{q^2 \xi_j}{(4\pi)^3} \,. \tag{A.3.3}$$

  In the collinear limit, we have $y_{ij} \to 1$, which gives the collinear limit of eq. (A.1.20) as

$$\lim_{y_{ij} \to 1} = \frac{q^2 - M_{\text{rec}}^2}{2q^0} - k_j^0 = \bar{k}_i^0 - \frac{q^0}{2} \xi_j \,, \tag{A.3.4}$$

  and thus

$$\lim_{y_{ij} \to 1} \mathcal{J}^{(\alpha)} = 1 - \frac{q^0}{2} \frac{\xi_j}{\bar{k}_i^0} \,. \tag{A.3.5}$$

- FSR, $m_i \neq 0$, $i > n_I \mid (i,j) \in \mathcal{P}_{\text{FKS}}$:

  In the case of FSR with a massless emitter $\mathcal{I}_i$, we have substituted y_{ij} with z, as seen in appendix A.1.1. We then have [68]

$$d\Phi_{\text{rad}}^{(\alpha)} = \mathcal{J}^{(\alpha)} d\xi_j dz d\phi_j \,, \quad \mathcal{J}^{(\alpha)} = \frac{q^2}{(2\pi)^3} \frac{k_j^0}{4\sqrt{\bar{k}_i^2}} \,. \tag{A.3.6}$$

- ISR, $m_i = 0$, $i \leq n_I \mid (i,j) \in \mathcal{P}_{\text{FKS}}$:

  For ISR, we have to account for the change of variables in eq. (A.2.11), where it is easy to see that we have

$$dx_\oplus dx_\ominus = \frac{d\bar{x}_\oplus d\bar{x}_\ominus}{1 - \xi_j} \,. \tag{A.3.7}$$

The Jacobian $\mathcal{J}^{(\alpha)}$ then is simply given by [67]

$$\mathcal{J}^{(\alpha)} = \frac{q^2}{(4\pi)^3} \frac{\xi_j}{1-\xi_j} \,. \tag{A.3.8}$$

B. Rescaling the Upper Integration Bound of ξ

Let us review the integration over ξ in eq. (2.2.20). There, we conveniently ignored that in general the upper bound $\xi_{\max}(y) \leq 1$ of the integration can be dependent on y, like in the case of massive final-state emitters (see eq. (A.1.45)) or ISR (see eq. (A.2.15)). This dependence makes the integration over ξ in eq. (2.2.20) more involved. However, by transforming ξ in the integration as [59]

$$\xi = \xi_{\max}(y)\tilde{\xi}\,, \tag{B.0.1}$$

we get the integral over $\tilde{\xi} \in [0,1]$ instead. To properly understand the effect of this rescaling, we have to reevaluate the integral of eq. (2.2.6) for $\tilde{\xi}$. We assume $\xi_{\mathrm{c}} = 1$ for the sake of simplicity (i.e. we use the unmodified plus distributions) and then have

$$\begin{aligned}
I &= \int_0^1 d\tilde{\xi}\xi_{\max}(y)\left(\tilde{\xi}\xi_{\max}(y)\right)^{-1-2\epsilon} f(\xi,y) && \text{(B.0.2)}\\
&= \int_0^1 d\tilde{\xi}\xi_{\max}(y)\frac{f(\xi,y)-f(0,y)}{\left(\tilde{\xi}\xi_{\max}(y)\right)^{1+2\epsilon}} + f(0,y)\int_0^1 d\tilde{\xi}\tilde{\xi}^{-1-2\epsilon}\xi_{\max}^{-2\epsilon}(y) && \text{(B.0.3)}\\
&= \int_0^1 d\tilde{\xi}\left[\left(\frac{1}{\tilde{\xi}}\right)_+ - 2\epsilon\left(\frac{\log\left(\tilde{\xi}\xi_{\max}(y)\right)}{\tilde{\xi}}\right)_+ + \mathcal{O}(\epsilon^2)\right] f(\xi,y) && \text{(B.0.4)}\\
&\quad + f(0,y)\left(-\frac{1}{2\epsilon} + \log\xi_{\max}(y) - \epsilon\log^2\xi_{\max}(y) + \mathcal{O}(\epsilon^2)\right),
\end{aligned}$$

where we expanded $\xi_{\max}^{-2\epsilon} = e^{-2\epsilon\log\xi_{\max}}$ in the last line. By rearranging terms we can write

$$\begin{aligned}
I &= \int_0^1 d\tilde{\xi}\left[\left(\frac{1}{\tilde{\xi}}\right)_+ + \log\xi_{\max}(y)\delta(\tilde{\xi})\right] f(\xi,y)\\
&\quad - 2\epsilon\int_0^1 d\tilde{\xi}\left[\left(\frac{\log\left(\tilde{\xi}\xi_{\max}(y)\right)}{\tilde{\xi}}\right)_+ + \frac{1}{2}\log^2\xi_{\max}(y)\delta(\tilde{\xi})\right] f(\xi,y) && \text{(B.0.5)}\\
&\quad + \int_0^1 d\tilde{\xi}\left(-\frac{1}{2\epsilon}\delta(\tilde{\xi}) + \mathcal{O}(\epsilon^2)\right) f(\xi,y)\,.
\end{aligned}$$

The content of the square brackets can thus be interpreted as the rescaled counterparts of the modified plus distributions in eq. (2.2.10) for $\xi_c = 1$. This directly leads to additional terms in in the last two lines of eq. (2.2.17), i.e. in the integrals $I_{+\delta,\alpha}$ and $I_{++,\alpha}$. For $I_{+\delta,\alpha}$, we already rescale the integration in eq. (2.3.26) in terms of z in the integration performed in section 2.3.2 and take care of the resulting terms. What remains are finite contributions to the integral $I_{++,\alpha}$ of the subtracted real contribution. Setting $\delta_{I/O} = 2$ in order to simplify the equation, in the case of FSR eq. (2.2.20) becomes

$$I_{++,\alpha} = \int_0^{2\pi} d\phi_j \int_{-1}^{1} \frac{dy_{ij}}{1-y_{ij}} \Bigg\{ \Big[\log \xi_{\max}(y_{ij})\hat{\mathcal{R}}_{ij}(0,y_{ij}) - \log \xi_{\max}(1)\hat{\mathcal{R}}_{ij}(0,1)\Big] + \int_0^1 \frac{d\tilde{\xi}_j}{\tilde{\xi}_j} \Big[\hat{\mathcal{R}}_{ij}(\tilde{\xi}_j \xi_{\max}(y_{ij}), y_{ij}) - \hat{\mathcal{R}}_{ij}(0,y_{ij}) - \hat{\mathcal{R}}_{ij}(\tilde{\xi}_j \xi_{\max}(1), 1) + \hat{\mathcal{R}}_{ij}(0,1)\Big] \Bigg\} . \tag{B.0.6}$$

Besides the integration over $\tilde{\xi}$ instead of ξ in the second row, the main difference to eq. (2.2.20) is the addition of

$$R_{\mathrm{FSR}}(y_{ij}) \equiv \log \xi_{\max}(y_{ij})\hat{\mathcal{R}}_{ij}(0,y_{ij}) - \log \xi_{\max}(1)\hat{\mathcal{R}}_{ij}(0,1) \tag{B.0.7}$$

in the first row, which we call the *soft remnant* for FSR. In the case of ISR, an additional term appears in the soft remnant due to collinear divergences for $y_j = -1$ in addition to $y = 1$ (see eq. (2.2.73)). By using $\xi_{\max}(y_j)$ for the case of ISR as given in eq. (A.2.15), we find $\xi_{\max}(\pm 1) = 1 - x_{\oplus}$ in the collinear limits, with the parton momentum fractions $x_{\oplus}$ as defined in eq. (A.2.11). The soft remnant in the case of ISR is then given by

$$R_{\mathrm{ISR}}(y_j) \equiv \log \xi_{\max}(y_j)\hat{\mathcal{R}}_{ij}(0,y_j) - \log(1-x_{\oplus})\hat{\mathcal{R}}_{ij}(0,1) - \log(1-x_{\ominus})\hat{\mathcal{R}}_{ij}(0,-1) . \tag{B.0.8}$$

C. Altarelli-Parisi Splitting Functions

C.1. Polarized Splitting Functions

The unregularized Altarelli-Parisi splitting functions $\hat{P}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}(z,k_\perp;\epsilon)$ in $d=4-2\epsilon$ dimensions (see [103]) can be brought into their polarized form $\hat{P}^{\lambda}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}(z,k_\perp;\epsilon)$. Here, λ denotes the pair of spin indices of the parent parton $\bar{\mathcal{I}}_i$, which are either s_i, s_i' if $\bar{\mathcal{I}}_i$ is a fermion or the Lorentz indices μ,ν if it is a gluon and we define [40]

$$\hat{P}^{s_i s_i'}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}(z,k_\perp;\epsilon) \equiv \left\langle s_i \middle| \hat{P}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}(z,k_\perp;\epsilon) \middle| s_i' \right\rangle , \tag{C.1.1}$$

$$\hat{P}^{\mu\nu}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}(z,k_\perp;\epsilon) \equiv \left\langle \mu \middle| \hat{P}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}(z,k_\perp;\epsilon) \middle| \nu \right\rangle . \tag{C.1.2}$$

With the parametrization for the four-momentum $\bar{k}_i$ of the emitter as in eq. (2.2.23), we then have [40, 59, 103]

$$\hat{P}^{s_i s_i'}_{q\to qg}(z,k_\perp;\epsilon) = \hat{P}^{s_i s_i'}_{\bar{q}\to\bar{q}g}(z,k_\perp;\epsilon) = \delta^{s_i s_i'} C_F \left[\frac{1+(1-z)^2}{z} - \epsilon z\right] , \tag{C.1.3}$$

$$\hat{P}^{s_i s_i'}_{q\to gq}(z,k_\perp;\epsilon) = \hat{P}^{s_i s_i'}_{\bar{q}\to g\bar{q}}(z,k_\perp;\epsilon) = \delta^{s_i s_i'} C_F \left[\frac{1+z^2}{1-z} - \epsilon\,(1-z)\right] , \tag{C.1.4}$$

$$\hat{P}^{\mu\nu}_{g\to q\bar{q}}(z,k_\perp;\epsilon) = T_F\,(1-\epsilon)\left[-g^{\mu\nu} - 4z\,(1-z)\,\frac{k_\perp^\mu k_\perp^\nu}{k_\perp^2}\right] , \tag{C.1.5}$$

$$\hat{P}^{\mu\nu}_{g\to gg}(z,k_\perp;\epsilon) = C_A\left[-2\left(\frac{z}{1-z}+\frac{1-z}{z}\right)g^{\mu\nu} + 4z\,(1-\epsilon)\,(1-z)\,\frac{k_\perp^\mu k_\perp^\nu}{k_\perp^2}\right] . \tag{C.1.6}$$

We note that eqs. (C.1.3) and (C.1.4) are symmetric under the change $z\to 1-z$. Moreover, we note that because of the fact that we used a slightly different ansatz following [59] for the parametrization in eq. (2.2.23), eqs. (C.1.5) and (C.1.6) have a different sign for the second term in the square brackets, when compared to other literature as e.g. [40, 105].

C.2. Spin-averaged Splitting Functions

In sections 2.3.2 and 2.3.3, we have used the spin-averaged splitting functions $\langle\hat{P}_{\bar{\mathcal{I}}_i\to\mathcal{I}_i\mathcal{I}_j}\rangle$. This is because of the angular integration in $d=4-2\epsilon$ dimensions over the angular

phase-space element $d\Omega^{d-2}$ in the collinear integrals in eqs. (2.3.22) and (2.3.39). While the splitting functions for the quark in eqs. (C.1.3) and (C.1.4) are obviously diagonal in the spin indices due to the fact that the quark-gluon coupling conserves helicity, the splitting functions of the gluon in eqs. (C.1.5) and (C.1.6) become diagonal after the phase-space integration. The integration requires the azimuthal average over the $d-2$ transverse components, given by [213]

$$-\left\langle \frac{k_\perp^\mu k_\perp^\nu}{k_\perp^2} \right\rangle_\phi = \frac{1}{d-2}\left(-g^{\mu\nu} + \frac{\bar{k}_i^\mu \eta^\nu + \eta^\mu \bar{k}_i^\nu}{\bar{k}\cdot\eta}\right) . \tag{C.2.1}$$

Here, η is defined as in eq. (2.2.23), i.e. $\eta \cdot k_\perp = 0$. Thus, by multiplying eq. (C.2.1) with eqs. (C.1.5) and (C.1.6), we are left with the spin-averaged splitting functions [40, 103]

$$\langle \hat{P}_{q\to qg} \rangle(z,\epsilon) = \langle \hat{P}_{\bar{q}\to \bar{q}g} \rangle(z,\epsilon) = C_F \left[\frac{1+(1-z)^2}{z} - \epsilon z \right] , \tag{C.2.2}$$

$$\langle \hat{P}_{q\to gq} \rangle(z,\epsilon) = \langle \hat{P}_{\bar{q}\to g\bar{q}} \rangle(z,\epsilon) = C_F \left[\frac{1+z^2}{1-z} - \epsilon\,(1-z) \right] , \tag{C.2.3}$$

$$\langle \hat{P}_{g\to q\bar{q}} \rangle(z,\epsilon) = T_F \left[1 - \frac{2z\,(1-z)}{1-\epsilon} \right] , \tag{C.2.4}$$

$$\langle \hat{P}_{g\to gg} \rangle(z,\epsilon) = 2C_A \left[\frac{z}{1-z} + \frac{1-z}{z} + z\,(1-z) \right] . \tag{C.2.5}$$

D. Supplements to the Virtual-subtracted Contributions

D.1. The Squared Massive Bare One-loop Amplitude

In eq. (2.3.1) we have given the squared bare one-loop amplitude $\mathcal{V}_b$ for the pure massless case. Here, we will give the explicit form of the more general expression

$$\mathcal{V}_b(\mathbf{\Phi}_n, f_B) = \mathcal{D}(\epsilon)\frac{\alpha_S}{2\pi}\Big[\mathcal{V}_m\Big]_{f_B}^{\mathbf{\Phi}_n}, \tag{D.1.1}$$

where the subscript of $\mathcal{V}_m$ denotes the inclusion of massive particles. Like in eq. (2.3.1), we denote by the superscript $\mathbf{\Phi}_n$ and the subscript f_B for the outer brackets the fact that all quantities inside of the outer brackets are computed in the n-body phase space for a Born flavor structure f_B. The terms $\mathcal{V}_m$ in the square bracket are given in [42] as

$$\begin{aligned}
\mathcal{V}_m = &-\left[\sum_{k=n_F}^{n_L^{(B)}+2}\left(\frac{1}{\epsilon^2}C(\bar{\mathcal{I}}_k)+\frac{1}{\epsilon}\gamma(\bar{\mathcal{I}}_k)\right)+\frac{1}{\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H+2}C(\bar{\mathcal{I}}_k)\right]\mathcal{B}\\
&+\frac{1}{\epsilon}\sum_{k=n_F}^{n_L^{(B)}+2}\sum_{k=k+1}^{n_L^{(B)}+n_H+2}\log\frac{2\bar{k}_k\cdot\bar{k}_l}{Q^2}\mathcal{B}_{kl}\\
&+\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H+1}\sum_{k=k+1}^{n_L^{(B)}+n_H+2}\frac{1}{v_{kl}}\log\frac{1+v_{kl}}{1-v_{kl}}\mathcal{B}_{kl}\\
&-\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H+2}\log\frac{m_k^2}{Q^2}\sum_{l=n_F}^{n_L^{(B)}+2}\mathcal{B}_{kl}+\mathcal{V}_{\text{fin}}\,,
\end{aligned} \tag{D.1.2}$$

with n_i, $n_L^{(B)}$, and n_H as defined in eq. (2.1.10). The quantities $C(\bar{\mathcal{I}})$ and $\gamma(\bar{\mathcal{I}})$ are defined in eqs. (2.2.46) and (2.2.47) and eq. (2.3.36), respectively. The term $\mathcal{V}_{\text{fin}}$ stands for the finite one-loop $\mathcal{O}(\epsilon^0)$-contributions, obtained by the evaluation of loop integrals. We define v_{kl} in eq. (D.2.9).

D.2. Massive Eikonal Integrals

In the following, we give an overview of the quantities $\hat{\mathcal{E}}_{kl}^{(m_k,m_l)}$ and $\mathcal{E}_{kl}^{(m_k,m_l)}$ that arise in the integration of the eikonal factors, given in eq. (2.3.9). They thus contribute to the result for the integrated soft subtraction term in eq. (2.3.18) and their massless forms (i.e. $m_k = m_l = 0$) have already been given in eqs. (2.3.19) and (2.3.20). The massive-massless and massive cases are dealt with in [68] and [100], with the analytic results of the latter published in [42], from which we cite in the following. As usual, $\mathcal{D}(\epsilon)$ is the normalization factor given in eq. (2.3.2), where we also introduced Q as the Ellis-Sexton scale. We note that in the soft limit the barred momenta of the set $\bar{\mathbf{\Phi}}_n$ are equal to their counterpart in $\mathbf{\Phi}_{n+1}$.

- $m_k = 0$, $m_l \neq 0$

$$\hat{\mathcal{E}}_{kl}^{(0,m_l)} = \mathcal{D}(\epsilon)\left[\frac{1}{2\epsilon^2} - \frac{1}{\epsilon}\left(\log\frac{2\bar{k}_k\cdot\bar{k}_l}{Q^2} + \frac{1}{2}\log\frac{4m_l^2\left(\bar{k}_k^0\right)^2}{\xi_{\mathrm{c}}^2 s Q^2}\right)\right], \tag{D.2.1}$$

$$\begin{aligned}\mathcal{E}_{kl}^{(0,m_l)} = {}& \log\xi_{\mathrm{c}}\left(\log\frac{s\xi_{\mathrm{c}}}{Q^2} + 2\log\frac{\bar{k}_k\cdot\bar{k}_l}{m_l\bar{k}_k^0}\right) - \frac{\pi^2}{12} + \frac{1}{4}\log^2\frac{s}{Q^2} \\ & - \frac{1}{4}\log^2\frac{1+\beta_l}{1-\beta_l} + \frac{1}{2}\log^2\frac{\bar{k}_k\cdot\bar{k}_l}{(1-\beta_l)\,\bar{k}_k^0\bar{k}_l^0} + \log\frac{s}{Q^2}\log\frac{\bar{k}_k\cdot\bar{k}_l}{m_l\bar{k}_k^0} \\ & - \mathrm{Li}_2\left(1 - \frac{(1+\beta_l)\,\bar{k}_k^0\bar{k}_l^0}{\bar{k}_k\cdot\bar{k}_l}\right) + \mathrm{Li}_2\left(1 - \frac{\bar{k}_k\cdot\bar{k}_l}{(1-\beta_l)\,\bar{k}_k^0\bar{k}_l^0}\right).\end{aligned} \tag{D.2.2}$$

  with

$$\beta_l = \sqrt{1 - \frac{m_l^2}{\left(\bar{k}_l^0\right)^2}} \tag{D.2.3}$$

- $m_l \neq 0$, $k = l$

$$\hat{\mathcal{E}}_{ll}^{(m_l,m_l)} = \mathcal{D}(\epsilon)\left(-\frac{1}{\epsilon}\right), \tag{D.2.4}$$

$$\mathcal{E}_{ll}^{(m_l,m_l)} = \log\frac{s\xi_{\mathrm{c}}^2}{Q^2} - \frac{1}{\beta_l}\log\frac{1+\beta_l}{1-\beta_l}, \tag{D.2.5}$$

- $m_k \neq 0$, $m_l \neq 0$, $k \neq l$

$$\hat{\mathcal{E}}_{kl}^{(m_k,m_l)} = \mathcal{D}(\epsilon)\left(-\frac{1}{2\epsilon}\frac{1}{v_{kl}}\log\frac{1+v_{kl}}{1-v_{kl}}\right), \tag{D.2.6}$$

$$\begin{aligned}\mathcal{E}_{kl}^{(m_k,m_l)} = {}& \frac{1}{2v_{kl}}\log\frac{1+v_{kl}}{1-v_{kl}}\log\frac{s\xi_{\mathrm{c}}^2}{Q^2} \\ & + \frac{(1+v_{kl})\left(\bar{k}_k\cdot\bar{k}_l\right)^2}{2m_k^2}\left(J(\alpha_{kl}\bar{k}_k^0, \alpha_{kl}\bar{k}_k^0\beta_k) - J(\bar{k}_l^0, \bar{k}_l^0\beta_l)\right),\end{aligned} \tag{D.2.7}$$

where we define the function J as

$$J(x,y) \equiv \frac{1}{2\lambda\nu}\left[\log^2\frac{x-y}{x+y} + 4\mathrm{Li}_2\left(1-\frac{x+y}{\nu}\right) + 4\mathrm{Li}_2\left(1-\frac{x-y}{\nu}\right)\right], \quad \text{(D.2.8)}$$

with the other parameters defined as

$$v_{kl} \equiv \sqrt{1-\left(\frac{m_k m_l}{\bar{k}_k\cdot\bar{k}_l}\right)^2}, \quad \text{(D.2.9)}$$

$$\alpha_{kl} \equiv \frac{1+v_{kl}}{m_k^2}\bar{k}_k\cdot\bar{k}_l\,, \quad \text{(D.2.10)}$$

$$\lambda \equiv \alpha_{kl}\bar{k}_k^0 - \bar{k}_l^0\,, \quad \text{(D.2.11)}$$

$$\nu \equiv \frac{\alpha_{kl}^2 m_k^2 - m_l^2}{2\lambda}\,. \quad \text{(D.2.12)}$$

E. Singular-Region Data for $pp \to Zj$

We introduced the way WHIZARD handles the data for singular regions in section 3.1.3. In table E.1, we provide a further example of the region data stored in the `region_data_t` type of WHIZARD for the process $pp \to Zj$ at NLO QCD to demonstrate the ISR case. We restrict the proton- and jet-flavor content to $u, \bar{u}, g$ for simplicity, as for processes with colored initial and final states the number of singular α-regions `alr` becomes quite large. One feature that could not be shown in table 3.1 is the way we summarize regions where a gluon could become collinear to both initial-state partons, in which case we use the emitter index `em` $= 0$.

alr	f_r	i_{f_r}	em	mul	$\mathcal{P}_{\text{FKS}}$	f_B	i_{f_B}
1	[-2,-2,23,-2,-2]	1	2	2	(1,4),(1,5),(2,4),(2,5)	[-2,21,23,-2]	2
2	[-2,-2,23,-2,-2]	1	1	2	(1,4),(1,5),(2,4),(2,5)	[21,-2,23,-2]	5
3	[-2, 2,23,-2, 2]	2	4	1	(1,4),(2,5),(4,5)	[-2, 2,23,21]	1
4	[-2, 2,23,-2, 2]	2	2	1	(1,4),(2,5),(4,5)	[-2,21,23,-2]	2
5	[-2, 2,23, 2,-2]	3	1	1	(1,5),(2,4),(4,5)	[21, 2,23, 2]	6
6	[-2, 2,23,21,21]	4	4	1	(0,4),(0,5),(4,5)	[-2, 2,23,21]	1
7	[-2, 2,23,21,21]	4	0	2	(0,4),(0,5),(4,5)	[-2, 2,23,21]	1
8	[-2,21,23,21,-2]	5	2	1	(0,4),(2,5),(4,5)	[-2, 2,23,21]	1
9	[-2,21,23,-2,21]	6	4	1	(0,5),(2,4),(4,5)	[-2,21,23,-2]	2
10	[-2,21,23,-2,21]	6	0	1	(0,5),(2,4),(4,5)	[-2,21,23,-2]	2
11	[2,-2,23,-2, 2]	7	4	1	(1,5),(2,4),(4,5)	[2,-2,23,21]	3
12	[2,-2,23, 2,-2]	8	2	1	(1,4),(2,5),(4,5)	[2,21,23, 2]	4
13	[2,-2,23,-2, 2]	7	1	1	(1,5),(2,4),(4,5)	[21,-2,23,-2]	5
14	[2,-2,23,21,21]	9	4	1	(0,4),(0,5),(4,5)	[2,-2,23,21]	3
15	[2,-2,23,21,21]	9	0	2	(0,4),(0,5),(4,5)	[2,-2,23,21]	3
16	[2, 2,23, 2, 2]	10	2	2	(1,4),(1,5),(2,4),(2,5)	[2,21,23, 2]	4
17	[2, 2,23, 2, 2]	10	1	2	(1,4),(1,5),(2,4),(2,5)	[21, 2,23, 2]	6
18	[2,21,23,21, 2]	11	2	1	(0,4),(2,5),(4,5)	[2,-2,23,21]	3
19	[2,21,23, 2,21]	12	4	1	(0,5),(2,4),(4,5)	[2,21,23, 2]	4
20	[2,21,23, 2,21]	12	0	1	(0,5),(2,4),(4,5)	[2,21,23, 2]	4
21	[21,-2,23,21,-2]	13	1	1	(0,4),(1,5),(4,5)	[2,-2,23,21]	3
22	[21,-2,23,-2,21]	14	4	1	(0,5),(1,4),(4,5)	[21,-2,23,-2]	5
23	[21,-2,23,-2,21]	14	0	1	(0,5),(1,4),(4,5)	[21,-2,23,-2]	5
24	[21, 2,23,21, 2]	15	1	1	(0,4),(1,5),(4,5)	[-2, 2,23,21]	1
25	[21, 2,23, 2,21]	16	4	1	(0,5),(1,4),(4,5)	[21, 2,23, 2]	6
26	[21, 2,23, 2,21]	16	0	1	(0,5),(1,4),(4,5)	[21, 2,23, 2]	6
27	[21,21,23,-2, 2]	17	1	1	(1,4),(1,5),(2,4),(2,5)	[-2,21,23,-2]	2
28	[21,21,23, 2,-2]	18	1	1	(1,4),(1,5),(2,4),(2,5)	[2,21,23, 2]	4
29	[21,21,23,-2, 2]	17	2	1	(1,4),(1,5),(2,4),(2,5)	[21,-2,23,-2]	5
30	[21,21,23, 2,-2]	18	2	1	(1,4),(1,5),(2,4),(2,5)	[21, 2,23, 2]	6

Table E.1. The singular-region data stored in the `region_data_t` type of WHIZARD for the process $pp \to Zj$ at NLO QCD with proton- and jet-flavor content $u, \bar{u}, g$. The meaning of the columns and the interpretation of the table as a whole is described in section 3.1.3. Particles in the flavor structures $f_{r/B}$ are represented by PDG IDs [168].

F. Examples of Process Setups and Analyses

In the following, we will use the process $e^+e^- \to jjj$ to give give examples for the setup files and analyses used in the MC event-generator comparison in chapter 4 and for the calculation of the $e^+e^- \to jjjjj$ fNLO QCD cross section in chapter 5. The results obtained within these chapters will be reproducible with the help of the following examples for WHIZARD `3.0.0`, SHERPA `2.2.10`, and RIVET `2.7.2`. For MG5_aMC@NLO `2.7.3`, the run and parameter cards can be found at [177].

F.1. WHIZARD Setup for $e^+e^- \to jjj$

F.1.1. Setup for Cross-section Comparisons

An example of a SINDARIN file for the process $e^+e^- \to jjj$, with identical settings as for the fNLO QCD cross section comparison, is shown in listing F.1. The exact meaning of each setting can be found in [117, appendix A]. Most of them are self-explanatory, however, with comments in listing F.1 explaining the general meaning for each group of settings. The cuts and scale choice are explained in section 4.2.1, together with the list of other settings related to physics. For a comparable reproduction of the results in section 4.2.2, all settings except for the MCI and O'MEGA setup must be identical. A comparable reproduction of the results here means within typical numerical fluctuations with a significance as defined in eq. (4.2.8) of roughly $\sigma^{\text{sig}}_{\text{NLO}} \lesssim 2$. For different seeds, MCI choices, etc., this insignificant level of fluctuations can be expected. The VAMP equivalences were implemented for NLO after the results in this thesis had been produced and had thus been disabled explicitly in the SINDARIN file.

SINDARIN Settings for $e^+e^- \to jjj$ Integration

```
# General settings:

# EW settings
model = SM ("GF_MW_MZ")
$blha_ew_scheme = "alpha_qed"
GF = 1.16639E-5
```

```
# Masses and widths
wZ = 0.0
wtop = 0.0
wW = 0.0
wH = 0.0

mZ = 91.188
mW = 80.419002
mH = 125.0
ms = 0
mc = 0
mb = 0
mtop = 173.2
me = 0
mmu = 0
mtau = 1.777

# Set OpenLoops as matrix-element generator
# and use complex mass scheme
$method = "openloops"
?openloops_use_cms = true

# VAMP2 (MCI) and O'Mega settings needed for MPI
$integration_method = "vamp2"
$rng_method = "rng_stream"
?use_vamp_equivalences = false
?omega_openmp = false
openmp_num_threads = 1

# Settings for running of QCD coupling
?alphas_is_fixed = false
?alphas_from_mz = true
?alphas_from_lambda_qcd = false
alphas_nf = 5
alphas_order = 2
alphas = 0.118

# Order of EW and QCD couplings
alpha_power = 2
alphas_power = 1

# Jet definition
alias jet = u:U:d:D:s:S:c:C:b:B:gl
jet_algorithm = antikt_algorithm
jet_r = 0.5
```

```
# Gluon must not split into top pair
$exclude_gauge_splittings = "t"

# Start of process specific setup:

# CM energy
sqrts = 1 TeV

cuts = let subevt @clustered_jets = cluster [jet] in
       let subevt @pt_selected = select if Pt > 30 GeV [@clustered_jets] in
       let subevt @eta_selected = select if abs(Eta) < 4 [@pt_selected] in
       count [@eta_selected] >= 3

scale = let int njet = count [jet] in
        if njet == 3 then
          ( eval Pt [extract index 1 [jet]]
          + eval Pt [extract index 2 [jet]]
          + eval Pt [extract index 3 [jet]]) / 2
        elsif njet == 4 then
          ( eval Pt [extract index 1 [jet]]
          + eval Pt [extract index 2 [jet]]
          + eval Pt [extract index 3 [jet]]
          + eval Pt [extract index 4 [jet]]) / 2
        else
          sqrts
        endif

process jjj = e1, E1 => jet, jet, jet { nlo_calculation = full }
# Increase calls per iteration for real component
mult_call_real = 10
# Stop adaptation of grid after reaching rel. error goal
relative_error_goal = 0.009
# Start integration of defined process
integrate (jjj) { iterations = 15:100000:"gw", 5:100000}
```

Listing F.1 Example for a SINDARIN file for the process $e^+e^- \to jjj$ as used in the calculation of the fNLO cross section comparison in section 4.2.

F.1.2. Convergence of Integration

With the setup as in listing F.1, the output of WHIZARD from the integration of the real component will be comparable to listing F.2, where we reduced the number of iterations in the grid adaptation from 15 to 10 for readability. As can be seen by

the iterative reduction of the relative error `Err[%]` and accuracy `Acc` towards a stable value the integration shows a convergent behaviour.

$e^+e^- \to jjj$ Real Component Integration

```
| Starting integration for process 'jjj' part 'real'
| Integrate: iterations = 10:1000000:"gw", 5:1000000
| Integrator: 1 chains, 2 channels, 8 dimensions
| Integrator: 1000000 initial calls, 20 max. bins, stratified = T
| Integrator: VAMP2
|=============================================================================|
| It      Calls  Integral[fb]  Error[fb]   Err[%]    Acc  Eff[%]   Chi2 N[It] |
|=============================================================================|
| VAMP2: Initialize new grids and write to file 'jjj.m2.vg2'.
| VAMP2: set chain: use chained weights.
   1     917504  4.0477851E+01  3.33E+00    8.24   78.91*   0.01
   2     917504  3.7674289E+01  7.78E-01    2.07   19.79*   0.18
   3     917504  3.7997520E+01  5.69E-01    1.50   14.35*   0.21
   4     917504  3.8023221E+01  6.58E-01    1.73   16.57    0.10
   5     917504  3.8612661E+01  6.71E-01    1.74   16.65    0.12
   6     917504  3.8380252E+01  6.01E-01    1.56   14.99*   0.22
   7     917504  3.8503573E+01  5.73E-01    1.49   14.25*   0.24
   8     917504  3.7414811E+01  5.24E-01    1.40   13.41*   0.41
   9     917504  3.8471833E+01  5.08E-01    1.32   12.64*   0.46
  10     917504  3.9402205E+01  5.25E-01    1.33   12.76    0.26
|-----------------------------------------------------------------------------|
  10    9175040  3.8321606E+01  1.95E-01    0.51   15.40    0.26    1.03  10
|-----------------------------------------------------------------------------|
  11     917504  3.8181039E+01  5.85E-01    1.53   14.68    0.16
  12     917504  3.9128814E+01  5.22E-01    1.33   12.77*   0.47
  13     917504  3.8743167E+01  5.05E-01    1.30   12.48*   0.58
  14     917504  3.8052397E+01  5.20E-01    1.37   13.10    0.50
  15     917504  3.8346431E+01  5.81E-01    1.51   14.51    0.15
|-----------------------------------------------------------------------------|
  15    4587520  3.8513061E+01  2.41E-01    0.63   13.42    0.15    0.70   5
|=============================================================================|
```

Listing F.2 Output from the integration of the real component for the process $e^+e^- \to jjj$.

F.1.3. Setup for Comparison of Differential Distributions

In order to reproduce the results from the comparison of differential distributions in section 4.3, we need to change a few settings regarding the process setup as described in section 4.3.1 and add the settings for event generation. These changes and additions to listing F.1 are shown in listing F.3.

Different/Additional Setup for $e^+e^- \to jjj$ Event Generation

```
# Differing process setup for fNLO event generation:

jet_algorithm = ee_genkt_algorithm
jet_p = -1
scale = mZ
?combined_nlo_integration = true
integrate (jjj) { iterations = 15:300000:"gw", 5:300000}

# Event generation setup:

?fixed_order_nlo_events = true
?unweighted = false
# fNLO (sub)events can have negative weights
?negative_weights = true
n_events=100000000
sample_format = hepmc
?hepmc_output_cross_section = true
$sample_normalization = "sigma"
# Do not write event groups where all subevents failed the cuts
?keep_failed_events = false
# Start event generation
simulate (jjj)
```

Listing F.3 Different and additional SINDARIN settings required in comparison to listing F.1 for the event generation as described in section 4.3.1.

F.2. Sherpa Run Card

An example of a SHERPA run card for the process $e^+e^- \to jjj$, with identical settings as for the fNLO QCD comparison of differential distributions, is shown in listing F.4. The exact meaning of each setting can be found in the SHERPA manual [214], with most of the settings being self-explanatory. In order to reproduce the results in section 4.3, the FASTJET interface of SHERPA has to be modified such that the option `antikt` of the `FastjetFinder` selector selects the generalized e^+e^- k_t jet algorithm instead as mentioned in section 4.3.1.

SHERPA run card for $e^+e^- \to jjj$

```
(run){
  % general settings
  EVENTS 1000000;
  ANALYSIS Rivet;

  % tags and settings for ME generators and integration
  LOOPGEN:=OpenLoops;
  ME_SIGNAL_GENERATOR Comix LOOPGEN;
  EVENT_GENERATION_MODE Weighted;
  VEGAS On;
  ERROR 0.005;

  % model parameters
  MODEL SM;
  EW_SCHEME 3;
  GF 1.16639E-5;
  ALPHAS(MZ) 0.118;
  ORDER_ALPHAS 1;
  MASSIVE[1] 0;
  MASSIVE[2] 0;
  MASSIVE[3] 0;
  MASSIVE[4] 0;
  MASSIVE[5] 0;
  MASSIVE[6] 1;
  MASS[1] 0.0;
  MASS[2] 0.0;
  MASS[3] 0.0;
  MASS[4] 0.0;
  MASS[5] 0.0;
  MASS[6] 173.2;
  MASSIVE[11] 0;
  MASSIVE[13] 0;
```

```
  MASSIVE[15] 1;
  MASS[11] 0.0;
  MASS[13] 0.0;
  MASS[15] 1.777;
  MASSIVE[23] 1;
  MASSIVE[24] 1;
  MASSIVE[25] 1;
  MASS[23] 91.188;
  MASS[24] 80.419002;
  MASS[25] 125.0;
  WIDTH[6] 0;
  WIDTH[15] 0;
  WIDTH[25] 0;
  WIDTH[24] 0;
  WIDTH[23] 0;
  STABLE[6] 1;
  STABLE[15] 1;
  STABLE[25] 1;
  STABLE[24] 1;
  STABLE[23] 1;

  % No shower, hadronization, etc.
  WIDTHSCHEME CMS;
  FRAGMENTATION Off;
  MI_HANDLER None;
  SHOWER_GENERATOR None;
  DECAYMODEL Off;
  YFS_MODE 0;
  ME_QED Off;
  ME_QED_CLUSTERING Off;

  % collider setup
  BEAM_1  11; BEAM_ENERGY_1 500;
  BEAM_2 -11; BEAM_ENERGY_2 500;
  BEAM_REMNANTS Off;

  % Fixed scale
  SCALES VAR{sqr(91.188)};

}(run);

(isr){
  % No QED ISR
  PDF_LIBRARY None;
}(isr);
```

```
(selector){
  % Jet selection and cuts
  % Sherpa interface has to be modified for antikt -> ee_genkt
  FastjetFinder antikt 3 30. 0. 0.5 0.75 4.;
}(selector);

(processes){
  Process 11 -11 -> 93 93 93;
  NLO_QCD_Mode Fixed_Order;
  Order (*,2);
  Loop_Generator LOOPGEN;
  ME_Generator Comix;
  RS_ME_Generator Comix;
  INTEGRATION_ERROR 0.009;
  End process;
}(processes);

(analysis){
  % Interface to Rivet analysis
  BEGIN_RIVET {
    -a threejetpTR2;
    USE_HEPMC_SHORT 1;
    IGNOREBEAMS 1;
  } END_RIVET;
}(analysis);
```

Listing F.4 Example for a SHERPA run card for the process $e^+e^- \to jjj$ as used for the event generation and analysis as described in section 4.3.1.

F.3. Rivet Analysis

An example of a RIVET analysis for the process $e^+e^- \to jjj$, with identical settings as for the fNLO QCD comparison of differential distributions in section 4.3, is shown in listing F.5. This analysis is written for RIVET `2.7.2`, for the reasons stated in section 4.3.2, but can be converted to RIVET `3` in a few steps. For example, for RIVET `2` correlated fNLO event groups (see section 3.1.5) are treated by the externally provided `NLOHisto1D` class, whereas in RIVET `3` these event groups are natively treated by the standard histogram classes (see section 4.3.2).

RIVET Analysis for $e^+e^- \to jjj$

```
// -*- C++ -*-
#include "Rivet/Analysis.hh"
#include "Rivet/Projections/FinalState.hh"
#include "Rivet/Projections/FastJets.hh"
#include "Rivet/Projections/Thrust.hh"
#include "Rivet/Projections/ParisiTensor.hh"
#include "Rivet/Projections/Hemispheres.hh"

namespace Rivet {
  class threejetpTR2 : public Analysis {
#include "NLOHisto1D.cc"
    public:
      // Constructor
      DEFAULT_RIVET_ANALYSIS_CTOR(threejetpTR2);
      void init() {
        // Initialize projections
        const FinalState fs;
        declare(fs, "FS");
        FastJets jetfs(fs, FastJets::GENKTEE, 0.5);
        declare(jetfs, "Jets");
        const Thrust thrust(fs);
        declare(thrust, "Thrust");
        declare(ParisiTensor(fs), "Parisi");
        declare(Hemispheres(thrust), "Hemispheres");
        // Book histograms
        _h_pT[0] = bookNLOHisto1D("pT1", bwspace(15., 30., 500.0, 460.,
            150.));
        _h_pT[1] = bookNLOHisto1D("pT2", bwspace(15., 30., 490.0, 250.,
            150.));
        _h_pT[2] = bookNLOHisto1D("pT3", bwspace(15., 30., 333., 29., 80.));
        _h_eta[0] = bookNLOHisto1D("eta1", bwspace(10., 0., 3.50, 0., 0.5));
        _h_eta[1] = bookNLOHisto1D("eta2", bwspace(10., 0., 3.50, 0., 0.5));
```

```
    _h_eta[2] = bookNLOHisto1D("eta3", bwspace(10., 0., 3.50, 0., 0.7));
    _h_E[0] = bookNLOHisto1D("E1", bwspace(15, 32., 506., 486., 100.));
    _h_E[1] = bookNLOHisto1D("E2", bwspace(15, 30., 504., 440., 150.));
    _h_E[2] = bookNLOHisto1D("E3", bwspace(15, 30., 503., 30., 100.));
    _h_invMass[0] = bookNLOHisto1D("invMass", bwspace(15., 31., 970.,
        910., 200.));
    _h_thrust[0] = bookNLOHisto1D("thrust", 15, 0.003, 0.43);
    _h_thrust[1] = bookNLOHisto1D("thrustMajor", bwspace(15, 0.024, 0.71,
        0.2, 0.2));
    _h_thrust[2] = bookNLOHisto1D("oblateness", bwspace(15, 0., 0.71,
        0.05, 0.2));
    _h_parisi[0] = bookNLOHisto1D("parisi_C", 15, 0.01, 1.);
    _h_parisi[1] = bookNLOHisto1D("parisi_D", 14, -0.0049, 1.);
    _h_hemispheres[0] = bookNLOHisto1D("heavyJetmass", bwspace(15.,
        0.003, 0.42, 0.1, 0.1));
    _h_hemispheres[1] = bookNLOHisto1D("totalJetbroad", bwspace(15.,
        0.01, 0.41, 0.125, 0.3));
    _h_hemispheres[2] = bookNLOHisto1D("wideJetbroad", bwspace(15., 0.01,
        0.33, 0.125, 0.3));
  }
  // Perform the per-event analysis
  void analyze(const Event& event) {
    // Retrieve clustered jets, sorted by pT
    const Jets& jets = apply<FastJets>(event, "Jets").jetsByPt();
    // Define pT and eta cut
    const Cut& cuts = Cuts::pT > 30.*GeV && Cuts::abseta < 4.;
    // Apply cuts to clustered jets
    const Jets& jets_cut = filter_select(jets, cuts);
    // Select events with two or more jets fulfilling the cuts
    if (jets_cut.size() > 2 && event.weight() != 0) {
      // Apply event shape projections
      const Thrust& thrust = apply<Thrust>(event, "Thrust");
      const ParisiTensor& parisi = applyProjection<ParisiTensor>(event,
          "Parisi");
      const Hemispheres& hemisphere = applyProjection<Hemispheres>(event,
          "Hemispheres");
      // Fill histograms
      _h_thrust[0]->fill(1 - thrust.thrust(), event);
      _h_thrust[1]->fill(thrust.thrustMajor(), event);
      _h_thrust[2]->fill(thrust.oblateness(), event);
      _h_parisi[0]->fill(parisi.C(), event);
      _h_parisi[1]->fill(parisi.D(), event);
      _h_hemispheres[0]->fill(hemisphere.scaledM2high(), event);
      _h_hemispheres[1]->fill(hemisphere.Bsum(), event);
      _h_hemispheres[2]->fill(hemisphere.Bmax(), event);
```

```
      _h_invMass[0]->fill((jets[0].momentum() +
          jets[1].momentum()).mass(), event);
      _h_pT[0]->fill(jets[0].pT()/GeV, event);
      _h_pT[1]->fill(jets[1].pT()/GeV, event);
      _h_pT[2]->fill(jets[2].pT()/GeV, event);
      _h_eta[0]->fill(jets[0].abseta()/GeV, event);
      _h_eta[1]->fill(jets[1].abseta()/GeV, event);
      _h_eta[2]->fill(jets[2].abseta()/GeV, event);
      _h_E[0]->fill(jets[0].E()/GeV, event);
      _h_E[1]->fill(jets[1].E()/GeV, event);
      _h_E[2]->fill(jets[2].E()/GeV, event);
    }
    // Veto (sub)event if its weight is zero or if it failed the cuts
    else if (event.weight() == 0){
      vetoEvent;
    }
    else {
      vetoEvent;
    }
  }
  // Scale histograms by factor of cross section over sum of weights
  void finalize() {
    double scaleFac = crossSection()/sumOfWeights();
    for (const NLOHisto1DPtr &_h : _h_pT) {
      _h->finalize();
      scale(_h, scaleFac);
    }
    for (const NLOHisto1DPtr &_h : _h_E) {
      _h->finalize();
      scale(_h, scaleFac);
    }
    for (const NLOHisto1DPtr &_h : _h_eta) {
      _h->finalize();
      scale(_h, scaleFac);
    }
    for (const NLOHisto1DPtr &_h : _h_invMass) {
      _h->finalize();
      scale(_h, scaleFac);
    }
    for (const NLOHisto1DPtr &_h : _h_thrust) {
      _h->finalize();
      scale(_h, scaleFac);
    }
    for (const NLOHisto1DPtr &_h : _h_parisi) {
      _h->finalize();
```

```
        scale(_h, scaleFac);
      }
      for (const NLOHisto1DPtr &_h : _h_hemispheres) {
        _h->finalize();
        scale(_h, scaleFac);
      }
    }
    // Declare histogram pointers
    NLOHisto1DPtr _h_pT[3];
    NLOHisto1DPtr _h_E[3];
    NLOHisto1DPtr _h_eta[3];
    NLOHisto1DPtr _h_thrust[3];
    NLOHisto1DPtr _h_parisi[2];
    NLOHisto1DPtr _h_hemispheres[3];
    NLOHisto1DPtr _h_invMass[1];
  };
  DECLARE_RIVET_PLUGIN(threejetpTR2);
}
```

Listing F.5 RIVET analysis for the process $e^+e^- \to jjj$ used to obtain the results in section 4.3.

Bibliography

[1] F. Englert and R. Brout. *Broken Symmetry and the Mass of Gauge Vector Mesons*. Phys. Rev. Lett., **13** (1964). Ed. by J. C. Taylor, pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

[2] P. W. Higgs. *Broken symmetries, massless particles and gauge fields*. Phys. Lett., **12** (1964), pp. 132–133. DOI: 10.1016/0031-9163(64)91136-9.

[3] P. W. Higgs. *Broken Symmetries and the Masses of Gauge Bosons*. Phys. Rev. Lett., **13** (1964). Ed. by J. C. Taylor, pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[4] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. *Global Conservation Laws and Massless Particles*. Phys. Rev. Lett., **13** (1964). Ed. by J. C. Taylor, pp. 585–587. DOI: 10.1103/PhysRevLett.13.585.

[5] P. W. Higgs. *Spontaneous Symmetry Breakdown without Massless Bosons*. Phys. Rev., **145** (1966), pp. 1156–1163. DOI: 10.1103/PhysRev.145.1156.

[6] T. W. B. Kibble. *Symmetry Breaking in Non-Abelian Gauge Theories*. Phys. Rev., **155** (1967). Ed. by J. C. Taylor, pp. 1554–1561. DOI: 10.1103/PhysRev.155.1554.

[7] G. Aad et al. *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*. Phys. Lett. B, **716** (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[8] S. Chatrchyan et al. *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*. Phys. Lett. B, **716** (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[9] *LHC Machine*. JINST, **3** (2008). Ed. by L. Evans and P. Bryant, S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[10] R. Aaij et al. *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays*. Phys. Rev. Lett., **115** (2015), p. 072001. DOI: 10.1103/PhysRevLett.115.072001. arXiv: 1507.03414 [hep-ex].

[11] V. C. Rubin and W. K. Ford Jr. *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions.* Astrophys. J., **159** (1970), pp. 379–403. DOI: 10.1086/150317.

[12] K. G. Begeman, A. H. Broeils, and R. H. Sanders. *Extended rotation curves of spiral galaxies: Dark haloes and modified dynamics.* Mon. Not. Roy. Astron. Soc., **249** (1991), p. 523.

[13] D. Clowe, M. Bradac, A. H. Gonzalez, et al. *A direct empirical proof of the existence of dark matter.* Astrophys. J. Lett., **648** (2006), pp. L109–L113. DOI: 10.1086/508162. arXiv: astro-ph/0608407.

[14] J. A. Tyson, G. P. Kochanski, and I. P. Dell'Antonio. *Detailed mass map of CL0024+1654 from strong lensing.* Astrophys. J. Lett., **498** (1998), p. L107. DOI: 10.1086/311314. arXiv: astro-ph/9801193.

[15] A. Refregier. *Weak gravitational lensing by large scale structure.* Ann. Rev. Astron. Astrophys., **41** (2003), pp. 645–668. DOI: 10.1146/annurev.astro.41.111302.102207. arXiv: astro-ph/0307212.

[16] N. Aghanim et al. *Planck 2018 results. VI. Cosmological parameters.* Astron. Astrophys., **641** (2020), A6. DOI: 10.1051/0004-6361/201833910. arXiv: 1807.06209 [astro-ph.CO].

[17] S. Profumo, L. Giani, and O. F. Piattella. *An Introduction to Particle Dark Matter.* Universe, **5**.10 (2019), p. 213. DOI: 10.3390/universe5100213. arXiv: 1910.05610 [hep-ph].

[18] B. Abi et al. *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm.* Phys. Rev. Lett., **126**.14 (2021), p. 141801. DOI: 10.1103/PhysRevLett.126.141801. arXiv: 2104.03281 [hep-ex].

[19] *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report.* **10/2020** (2020). Ed. by I. Béjar Alonso, O. Brüning, P. Fessia, et al. DOI: 10.23731/CYRM-2020-0010.

[20] A. Abada et al. *HE-LHC: The High-Energy Large Hadron Collider: Future Circular Collider Conceptual Design Report Volume 4.* Eur. Phys. J. ST, **228**.5 (2019), pp. 1109–1382. DOI: 10.1140/epjst/e2019-900088-6.

[21] *The International Linear Collider Technical Design Report - Volume 1: Executive Summary* (2013). Ed. by T. Behnke, J. E. Brau, B. Foster, et al. arXiv: 1306.6327 [physics.acc-ph].

[22] R. Franceschini et al. *The CLIC Potential for New Physics.* **3/2018** (2018). Ed. by J. de Blas. DOI: 10.23731/CYRM-2018-003. arXiv: 1812.02093 [hep-ph].

[23] A. Abada et al. *FCC Physics Opportunities: Future Circular Collider Conceptual Design Report Volume 1.* Eur. Phys. J. C, **79**.6 (2019), p. 474. DOI: 10.1140/epjc/s10052-019-6904-3.

[24] A. Abada et al. *FCC-ee: The Lepton Collider: Future Circular Collider Conceptual Design Report Volume 2.* Eur. Phys. J. ST, **228**.2 (2019), pp. 261–623. DOI: 10.1140/epjst/e2019-900045-4.

[25] S. Marzani, G. Soyez, and M. Spannowsky. *Looking inside jets: an introduction to jet substructure and boosted-object phenomenology.* Vol. 958. Springer, 2019. DOI: 10.1007/978-3-030-15709-8. arXiv: 1901.10342 [hep-ph].

[26] C. Weiss. Top quark physics as a prime application of automated higher-order corrections. PhD thesis. Hamburg: Hamburg U., 2017. DOI: 10.3204/PUBDB-2017-07541.

[27] G. Dissertori, I. Knowles, and M. Schmelling. *Quantum Chromodynamics: High Energy Experiments and Theory.* International Series of Monographs on Physics. OUP Oxford, 2009. ISBN: 9780199566419.

[28] T. Kinoshita. *Mass singularities of Feynman amplitudes.* J. Math. Phys., **3** (1962), pp. 650–677. DOI: 10.1063/1.1724268.

[29] T. D. Lee and M. Nauenberg. *Degenerate Systems and Mass Singularities.* Phys. Rev., **133** (1964), B1549–B1562. DOI: 10.1103/PhysRev.133.B1549.

[30] F. Bloch and A. Nordsieck. *Note on the Radiation Field of the electron.* Phys. Rev., **52** (1937), pp. 54–59. DOI: 10.1103/PhysRev.52.54.

[31] A. Buckley et al. *General-purpose event generators for LHC physics.* Phys. Rept., **504** (2011), pp. 145–233. DOI: 10.1016/j.physrep.2011.03.005. arXiv: 1101.2599 [hep-ph].

[32] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer. *MadGraph 5 : Going Beyond.* JHEP, **06** (2011), p. 128. DOI: 10.1007/JHEP06(2011)128. arXiv: 1106.0522 [hep-ph].

[33] J. Alwall, R. Frederix, S. Frixione, et al. *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations.* JHEP, **07** (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[34] T. Gleisberg, S. Hoeche, F. Krauss, et al. *Event generation with SHERPA 1.1.* JHEP, **02** (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].

[35] W. Kilian, T. Ohl, and J. Reuter. *WHIZARD: Simulating Multi-Particle Processes at LHC and ILC*. Eur. Phys. J. C, **71** (2011), p. 1742. DOI: 10.1140/epjc/s10052-011-1742-y. arXiv: 0708.4233 [hep-ph].

[36] T. Sjostrand, S. Mrenna, and P. Z. Skands. *PYTHIA 6.4 Physics and Manual*. JHEP, **05** (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175.

[37] T. Sjöstrand, S. Ask, J. R. Christiansen, et al. *An introduction to PYTHIA 8.2*. Comput. Phys. Commun., **191** (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[38] M. Bahr et al. *Herwig++ Physics and Manual*. Eur. Phys. J. C, **58** (2008), pp. 639–707. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].

[39] J. Bellm et al. *Herwig 7.0/Herwig++ 3.0 release note*. Eur. Phys. J. C, **76**.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[40] S. Catani and M. Seymour. *A General algorithm for calculating jet cross-sections in NLO QCD*. Nucl. Phys. B, **485** (1997). [Erratum: Nucl.Phys.B 510, 503–504 (1998)], pp. 291–419. DOI: 10.1016/S0550-3213(96)00589-5. arXiv: hep-ph/9605323.

[41] S. Frixione, Z. Kunszt, and A. Signer. *Three jet cross-sections to next-to-leading order*. Nucl. Phys. B, **467** (1996), pp. 399–442. DOI: 10.1016/0550-3213(96)00110-1. arXiv: hep-ph/9512328.

[42] R. Frederix, S. Frixione, F. Maltoni, and T. Stelzer. *Automation of next-to-leading order computations in QCD: The FKS subtraction*. JHEP, **10** (2009), p. 003. DOI: 10.1088/1126-6708/2009/10/003. arXiv: 0908.4272 [hep-ph].

[43] R. K. Ellis and G. Zanderighi. *Scalar one-loop integrals for QCD*. JHEP, **02** (2008), p. 002. DOI: 10.1088/1126-6708/2008/02/002. arXiv: 0712.1851 [hep-ph].

[44] A. van Hameren. *OneLOop: For the evaluation of one-loop scalar functions*. Comput. Phys. Commun., **182** (2011), pp. 2427–2438. DOI: 10.1016/j.cpc.2011.06.011. arXiv: 1007.4716 [hep-ph].

[45] A. Denner, S. Dittmaier, and L. Hofer. *Collier: a fortran-based Complex One-Loop LIbrary in Extended Regularizations*. Comput. Phys. Commun., **212** (2017), pp. 220–238. DOI: 10.1016/j.cpc.2016.10.013. arXiv: 1604.06792 [hep-ph].

[46] G. Ossola, C. G. Papadopoulos, and R. Pittau. *CutTools: A Program implementing the OPP reduction method to compute one-loop amplitudes*. JHEP, **03** (2008), p. 042. DOI: 10.1088/1126-6708/2008/03/042. arXiv: 0711.3596 [hep-ph].

[47] A. van Hameren, C. Papadopoulos, and R. Pittau. *Automated one-loop calculations: A Proof of concept*. JHEP, **09** (2009), p. 106. DOI: 10.1088/1126-6708/2009/09/106. arXiv: 0903.4665 [hep-ph].

[48] G. Cullen et al. *GOSAM-2.0: a tool for automated one-loop calculations within the Standard Model and beyond*. Eur. Phys. J. C, **74**.8 (2014), p. 3001. DOI: 10.1140/epjc/s10052-014-3001-5. arXiv: 1404.7096 [hep-ph].

[49] V. Hirschi, R. Frederix, S. Frixione, et al. *Automation of one-loop QCD corrections*. JHEP, **05** (2011), p. 044. DOI: 10.1007/JHEP05(2011)044. arXiv: 1103.0621 [hep-ph].

[50] S. Badger, B. Biedermann, and P. Uwer. *NGluon: A Package to Calculate One-loop Multi-gluon Amplitudes*. Comput. Phys. Commun., **182** (2011), pp. 1674–1692. DOI: 10.1016/j.cpc.2011.04.008. arXiv: 1011.2900 [hep-ph].

[51] S. Badger, B. Biedermann, P. Uwer, and V. Yundin. *Numerical evaluation of virtual corrections to multi-jet production in massless QCD*. Comput. Phys. Commun., **184** (2013), pp. 1981–1998. DOI: 10.1016/j.cpc.2013.03.018. arXiv: 1209.0100 [hep-ph].

[52] F. Cascioli, P. Maierhofer, and S. Pozzorini. *Scattering Amplitudes with Open Loops*. Phys. Rev. Lett., **108** (2012), p. 111601. DOI: 10.1103/PhysRevLett.108.111601. arXiv: 1111.5206 [hep-ph].

[53] F. Buccioni, J.-N. Lang, J. M. Lindert, et al. *OpenLoops 2*. Eur. Phys. J. C, **79**.10 (2019), p. 866. DOI: 10.1140/epjc/s10052-019-7306-2. arXiv: 1907.13071 [hep-ph].

[54] S. Actis, A. Denner, L. Hofer, A. Scharf, and S. Uccirati. *Recursive generation of one-loop amplitudes in the Standard Model*. JHEP, **04** (2013), p. 037. DOI: 10.1007/JHEP04(2013)037. arXiv: 1211.6316 [hep-ph].

[55] S. Actis, A. Denner, L. Hofer, et al. *RECOLA: REcursive Computation of One-Loop Amplitudes*. Comput. Phys. Commun., **214** (2017), pp. 140–173. DOI: 10.1016/j.cpc.2017.01.004. arXiv: 1605.01090 [hep-ph].

[56] A. Denner, J.-N. Lang, and S. Uccirati. *NLO electroweak corrections in extended Higgs Sectors with RECOLA2*. JHEP, **07** (2017), p. 087. DOI: 10.1007/JHEP07(2017)087. arXiv: 1705.06053 [hep-ph].

[57] A. Denner, J.-N. Lang, and S. Uccirati. *Recola2: REcursive Computation of One-Loop Amplitudes 2*. Comput. Phys. Commun., **224** (2018), pp. 346–361. DOI: 10.1016/j.cpc.2017.11.013. arXiv: 1711.07388 [hep-ph].

[58] A. Kanaki and C. G. Papadopoulos. *HELAC: A Package to compute electroweak helicity amplitudes*. Comput. Phys. Commun., **132** (2000), pp. 306–315. DOI: 10.1016/S0010-4655(00)00151-X. arXiv: hep-ph/0002082.

[59] 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), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[60] K. Arnold et al. *VBFNLO: A Parton level Monte Carlo for processes with electroweak bosons*. Comput. Phys. Commun., **180** (2009), pp. 1661–1670. DOI: 10.1016/j.cpc.2009.03.006. arXiv: 0811.4559 [hep-ph].

[61] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, and G. Heinrich. *EERAD3: Event shapes and jet rates in electron-positron annihilation at order* α_s^3. Comput. Phys. Commun., **185** (2014), p. 3331. DOI: 10.1016/j.cpc.2014.07.024. arXiv: 1402.4140 [hep-ph].

[62] M. Grazzini, S. Kallweit, and M. Wiesemann. *Fully differential NNLO computations with MATRIX*. Eur. Phys. J. C, **78**.7 (2018), p. 537. DOI: 10.1140/epjc/s10052-018-5771-7. arXiv: 1711.06631 [hep-ph].

[63] R. Boughezal, J. M. Campbell, R. K. Ellis, et al. *Color singlet production at NNLO in MCFM*. Eur. Phys. J. C, **77**.1 (2017), p. 7. DOI: 10.1140/epjc/s10052-016-4558-y. arXiv: 1605.08011 [hep-ph].

[64] J. Currie, A. Gehrmann-De Ridder, T. Gehrmann, et al. *Jet cross sections at the LHC with NNLOJET*. PoS, **LL2018** (2018), p. 001. DOI: 10.22323/1.303.0001. arXiv: 1807.06057 [hep-ph].

[65] C. Anastasiou, C. Duhr, F. Dulat, et al. *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*. JHEP, **05** (2016), p. 058. DOI: 10.1007/JHEP05(2016)058. arXiv: 1602.00695 [hep-ph].

[66] S. Frixione. *A General approach to jet cross-sections in QCD*. Nucl. Phys. B, **507** (1997), pp. 295–314. DOI: 10.1016/S0550-3213(97)00574-9. arXiv: hep-ph/9706545.

[67] S. Frixione, P. Nason, and C. Oleari. *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*. JHEP, **11** (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[68] L. Barze, G. Montagna, P. Nason, O. Nicrosini, and F. Piccinini. *Implementation of electroweak corrections in the POWHEG BOX: single W production.* JHEP, **04** (2012), p. 037. DOI: 10.1007/JHEP04(2012)037. arXiv: 1202.0465 [hep-ph].

[69] T. Ježo and P. Nason. *On the Treatment of Resonances in Next-to-Leading Order Calculations Matched to a Parton Shower.* JHEP, **12** (2015), p. 065. DOI: 10.1007/JHEP12(2015)065. arXiv: 1509.09071 [hep-ph].

[70] R. Frederix, S. Frixione, V. Hirschi, et al. *The automation of next-to-leading order electroweak calculations.* JHEP, **07** (2018), p. 185. DOI: 10.1007/JHEP07(2018)185. arXiv: 1804.10017 [hep-ph].

[71] S. Frixione. *Initial conditions for electron and photon structure and fragmentation functions.* JHEP, **11** (2019), p. 158. DOI: 10.1007/JHEP11(2019)158. arXiv: 1909.03886 [hep-ph].

[72] V. Bertone, M. Cacciari, S. Frixione, and G. Stagnitto. *The partonic structure of the electron at the next-to-leading logarithmic accuracy in QED.* JHEP, **03** (2020), p. 135. DOI: 10.1007/JHEP03(2020)135. arXiv: 1911.12040 [hep-ph].

[73] B. Chokoufe Nejad. Scrutinizing the Top Quark at Lepton Colliders with Higher Orders: From Fixed Order to Resummation and Matching. PhD thesis. Hamburg: Hamburg U., 2017. DOI: 10.3204/PUBDB-2017-07624.

[74] G. F. Sterman and S. Weinberg. *Jets from Quantum Chromodynamics.* Phys. Rev. Lett., **39** (1977), p. 1436. DOI: 10.1103/PhysRevLett.39.1436.

[75] S. Brandt, C. Peyrou, R. Sosnowski, and A. Wroblewski. *The Principal axis of jets. An Attempt to analyze high-energy collisions as two-body processes.* Phys. Lett., **12** (1964), pp. 57–61. DOI: 10.1016/0031-9163(64)91176-X.

[76] E. Farhi. *Quantum Chromodynamics Test for Jets.* Phys. Rev. Lett., **39** (25 1977), pp. 1587–1588. DOI: 10.1103/PhysRevLett.39.1587.

[77] G. Parisi. *Super Inclusive Cross-Sections.* Phys. Lett. B, **74** (1978), pp. 65–67. DOI: 10.1016/0370-2693(78)90061-8.

[78] J. F. Donoghue, F. Low, and S.-Y. Pi. *Tensor Analysis of Hadronic Jets in Quantum Chromodynamics.* Phys. Rev. D, **20** (1979), p. 2759. DOI: 10.1103/PhysRevD.20.2759.

[79] R. K. Ellis, D. A. Ross, and A. E. Terrano. *The Perturbative Calculation of Jet Structure in e+ e- Annihilation.* Nucl. Phys. B, **178** (1981), pp. 421–456. DOI: 10.1016/0550-3213(81)90165-6.

[80] A. Pich. *Precision physics with inclusive QCD processes.* Prog. Part. Nucl. Phys., **117** (2021), p. 103846. DOI: 10.1016/j.ppnp.2020.103846. arXiv: 2012.04716 [hep-ph].

[81] A. Heister et al. *Measurements of the strong coupling constant and the QCD color factors using four jet observables from hadronic Z decays.* Eur. Phys. J. C, **27** (2003), pp. 1–17. DOI: 10.1140/epjc/s2002-01114-2.

[82] J. Abdallah et al. *Measurement of the energy dependence of hadronic jet rates and the strong coupling alpha(s) from the four-jet rate with the DELPHI detector at LEP.* Eur. Phys. J. C, **38** (2005), pp. 413–426. DOI: 10.1140/epjc/s2004-02060-7. arXiv: hep-ex/0410071.

[83] G. Abbiendi et al. *Measurement of the Strong Coupling alpha(s) from four-jet observables in e+ e- annihilation.* Eur. Phys. J. C, **47** (2006), pp. 295–307. DOI: 10.1140/epjc/s2006-02581-y. arXiv: hep-ex/0601048.

[84] R. Frederix, S. Frixione, K. Melnikov, and G. Zanderighi. *NLO QCD corrections to five-jet production at LEP and the extraction of $\alpha_s(M_Z)$.* JHEP, **11** (2010), p. 050. DOI: 10.1007/JHEP11(2010)050. arXiv: 1008.5313 [hep-ph].

[85] S. Catani, S. Dittmaier, M. H. Seymour, and Z. Trocsanyi. *The Dipole formalism for next-to-leading order QCD calculations with massive partons.* Nucl. Phys. B, **627** (2002), pp. 189–265. DOI: 10.1016/S0550-3213(02)00098-6. arXiv: hep-ph/0201036.

[86] T. Gleisberg and F. Krauss. *Automating dipole subtraction for QCD NLO calculations.* Eur. Phys. J. C, **53** (2008), pp. 501–523. DOI: 10.1140/epjc/s10052-007-0495-0. arXiv: 0709.2881 [hep-ph].

[87] C.-H. Chung and T. Robens. *Nagy-Soper subtraction scheme for multiparton final states.* Phys. Rev. D, **87** (2013), p. 074032. DOI: 10.1103/PhysRevD.87.074032. arXiv: 1209.1569 [hep-ph].

[88] M. Bach, C. H. Chung, and T. Robens. *An alternative subtraction scheme for NLO QCD calculations using Nagy-Soper dipoles.* PoS, **RADCOR2013** (2013), p. 014. DOI: 10.22323/1.197.0014. arXiv: 1311.5773 [hep-ph].

[89] D. A. Kosower. *Antenna factorization of gauge theory amplitudes.* Phys. Rev. D, **57** (1998), pp. 5410–5416. DOI: 10.1103/PhysRevD.57.5410. arXiv: hep-ph/9710213.

[90] J. M. Campbell, M. A. Cullen, and E. W. N. Glover. *Four jet event shapes in electron - positron annihilation.* Eur. Phys. J. C, **9** (1999), pp. 245–265. DOI: 10.1007/s100529900034. arXiv: hep-ph/9809429.

[91] D. A. Kosower. *Antenna factorization in strongly ordered limits.* Phys. Rev. D, **71** (2005), p. 045016. DOI: 10.1103/PhysRevD.71.045016. arXiv: hep-ph/0311272.

[92] Z. Nagy and D. E. Soper. *Parton showers with quantum interference*. JHEP, **09** (2007), p. 114. DOI: 10.1088/1126-6708/2007/09/114. arXiv: 0706.0017 [hep-ph].

[93] Z. Nagy and D. E. Soper. *Parton showers with quantum interference: Leading color, spin averaged*. JHEP, **03** (2008), p. 030. DOI: 10.1088/1126-6708/2008/03/030. arXiv: 0801.1917 [hep-ph].

[94] Z. Nagy and D. E. Soper. *Parton showers with quantum interference: Leading color, with spin*. JHEP, **07** (2008), p. 025. DOI: 10.1088/1126-6708/2008/07/025. arXiv: 0805.0216 [hep-ph].

[95] T. Robens. *Nagy-Soper Subtraction: A Review*. Mod. Phys. Lett. A, **28** (2013), p. 1330020. DOI: 10.1142/S0217732313300206. arXiv: 1306.1946 [hep-ph].

[96] A. Gehrmann-De Ridder, T. Gehrmann, and E. W. N. Glover. *Antenna subtraction at NNLO*. JHEP, **09** (2005), p. 056. DOI: 10.1088/1126-6708/2005/09/056. arXiv: hep-ph/0505111.

[97] C. Weiss, B. Chokoufe Nejad, W. Kilian, and J. Reuter. *Automated NLO QCD Corrections with WHIZARD*. PoS, **EPS-HEP2015** (2015), p. 466. DOI: 10.22323/1.234.0466. arXiv: 1510.02666 [hep-ph].

[98] P. Stienemeier, S. Braß, P. Bredt, et al. "WHIZARD 3.0: Status and News". *International Workshop on Future Linear Colliders*. 2021. arXiv: 2104.11141 [hep-ph].

[99] F. Bach, B. C. Nejad, A. Hoang, et al. *Fully-differential Top-Pair Production at a Lepton Collider: From Threshold to Continuum*. JHEP, **03** (2018), p. 184. DOI: 10.1007/JHEP03(2018)184. arXiv: 1712.02220 [hep-ph].

[100] S. Frixione, E. Laenen, P. Motylinski, and B. R. Webber. *Single-top production in MC@NLO*. JHEP, **03** (2006), p. 092. DOI: 10.1088/1126-6708/2006/03/092. arXiv: hep-ph/0512250.

[101] M. E. Peskin and D. V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995. ISBN: 978-0-201-50397-5.

[102] V. Sudakov. *Vertex parts at very high-energies in quantum electrodynamics*. Sov. Phys. JETP, **3** (1956), pp. 65–71.

[103] G. Altarelli and G. Parisi. *Asymptotic Freedom in Parton Language*. Nucl. Phys. B, **126** (1977), pp. 298–318. DOI: 10.1016/0550-3213(77)90384-4.

[104] H. Murayama, I. Watanabe, and K. Hagiwara. *HELAS: HELicity amplitude subroutines for Feynman diagram evaluations* (1992).

[105] S. Catani and M. Grazzini. *Collinear factorization and splitting functions for next-to-next-to-leading order QCD calculations*. Phys. Lett. B, **446** (1999), pp. 143–152. DOI: 10.1016/S0370-2693(98)01513-5. arXiv: hep-ph/9810389.

[106] Y. L. Dokshitzer, V. A. Khoze, A. H. Mueller, and S. Troian. *Basics of perturbative QCD*. Editions Frontières, 1991. ISBN: 9782863321010.

[107] S. Catani and M. Grazzini. *The soft gluon current at one loop order*. Nucl. Phys. B, **591** (2000), pp. 435–454. DOI: 10.1016/S0550-3213(00)00572-1. arXiv: hep-ph/0007142.

[108] A. Bassetto, M. Ciafaloni, and G. Marchesini. *Jet Structure and Infrared Sensitive Quantities in Perturbative QCD*. Phys. Rept., **100** (1983), pp. 201–272. DOI: 10.1016/0370-1573(83)90083-2.

[109] A. V. Efremov and A. V. Radyushkin. *Interacting Partons*. Phys. Lett. B, **63** (1976). Ed. by V. I. Zhuravlev, p. 449. DOI: 10.1016/0370-2693(76)90394-4.

[110] H. D. Politzer. *Gluon Corrections to Drell-Yan Processes*. Nucl. Phys. B, **129** (1977), pp. 301–318. DOI: 10.1016/0550-3213(77)90197-3.

[111] J. C. Collins, D. E. Soper, and G. F. Sterman. Factorization of Hard Processes in QCD. Vol. 5. 1989, pp. 1–91. DOI: 10.1142/9789814503266_0001. arXiv: hep-ph/0409313.

[112] R. K. Ellis, W. J. Stirling, and B. R. Webber. *QCD and collider physics*. Vol. 8. Cambridge University Press, 2011. ISBN: 978-0-521-54589-1.

[113] B. Potter. *Inclusive single- and dijet rates in next-to-leading order QCD for $\gamma^* p$ and $\gamma^*\gamma$ collisions*. Eur. Phys. J. direct, **1**.1 (1999), p. 5. DOI: 10.1007/s1010599c0005. arXiv: hep-ph/9707319.

[114] R. Ellis and J. Sexton. *QCD Radiative Corrections to Parton Parton Scattering*. Nucl. Phys. B, **269** (1986), pp. 445–484. DOI: 10.1016/0550-3213(86)90232-4.

[115] Z. Kunszt and D. E. Soper. *Calculation of jet cross-sections in hadron collisions at order α_S^3*. Phys. Rev. D, **46** (1992), pp. 192–221. DOI: 10.1103/PhysRevD.46.192.

[116] Z. Kunszt, A. Signer, and Z. Trocsanyi. *Singular terms of helicity amplitudes at one loop in QCD and the soft limit of the cross-sections of multiparton processes*. Nucl. Phys. B, **420** (1994), pp. 550–564. DOI: 10.1016/0550-3213(94)90077-9. arXiv: hep-ph/9401294.

[117] W. Kilian, T. Ohl, and J. Reuter. *The WHIZARD Event Generator*. https://whizard.hepforge.org/manual.pdf. [Online; accessed 13.06.2021].

[118] J. Reuter, F. Bach, B. Chokoufe, et al. *Modern Particle Physics Event Generation with WHIZARD*. J. Phys. Conf. Ser., **608**.1 (2015). Ed. by L. Fiala, M. Lokajicek, and N. Tumova, p. 012063. DOI: 10.1088/1742-6596/608/1/012063. arXiv: 1410.4505 [hep-ph].

[119] W. Kilian, S. Braß, T. Ohl, et al. "New Developments in WHIZARD Version 2.6". *International Workshop on Future Linear Collider*. 2018. arXiv: 1801.08034 [hep-ph].

[120] J. Reuter, S. Braß, P. Bredt, et al. "Status of the WHIZARD generator for linear colliders". *International Workshop on Future Linear Colliders*. 2020. arXiv: 2002.06122 [hep-ph].

[121] J. Aguilar-Saavedra et al. *TESLA: The Superconducting electron positron linear collider with an integrated x-ray laser laboratory. Technical design report. Part 3. Physics at an e+ e- linear collider* (2001). arXiv: hep-ph/0106315.

[122] R. Chierici, S. Rosati, and M. Kobel. *Strong electroweak symmetry breaking signals in w w scattering at tesla*. AIP Conf. Proc., **578**.1 (2001). Ed. by T. Behnke, S. Bertolucci, R. D. Heuer, et al., pp. 544–549. DOI: 10.1063/1.1394378.

[123] *The International Linear Collider Technical Design Report - Volume 2: Physics* (2013). Ed. by H. Baer et al. arXiv: 1306.6352 [hep-ph].

[124] H. Abramowicz et al. *The International Linear Collider Technical Design Report - Volume 4: Detectors* (2013). Ed. by T. Behnke, J. E. Brau, P. N. Burrows, et al. arXiv: 1306.6329 [physics.ins-det].

[125] K. Fujii et al. *Physics Case for the International Linear Collider* (2015). arXiv: 1506.05992 [hep-ex].

[126] M. Moretti, T. Ohl, and J. Reuter. *O'Mega: An Optimizing matrix element generator* (2001). Ed. by T. Behnke, S. Bertolucci, R. D. Heuer, et al., pp. 1981–2009. arXiv: hep-ph/0102195.

[127] W. Kilian, T. Ohl, J. Reuter, and C. Speckner. *QCD in the Color-Flow Representation*. JHEP, **10** (2012), p. 022. DOI: 10.1007/JHEP10(2012)022. arXiv: 1206.3700 [hep-ph].

[128] B. Chokoufe Nejad, T. Ohl, and J. Reuter. *Simple, parallel virtual machines for extreme computations*. Comput. Phys. Commun., **196** (2015), pp. 58–69. DOI: 10.1016/j.cpc.2015.05.015. arXiv: 1411.3834 [physics.comp-ph].

[129] T. Binoth et al. *A Proposal for a Standard Interface between Monte Carlo Tools and One-Loop Programs*. Comput. Phys. Commun., **181** (2010). Ed. by G. Belanger, F. Boudjema, and J.-P. Guillet, pp. 1612–1622. DOI: 10.1016/j.cpc.2010.05.016. arXiv: 1001.1307 [hep-ph].

[130] S. Alioli et al. *Update of the Binoth Les Houches Accord for a standard interface between Monte Carlo tools and one-loop programs*. Comput. Phys. Commun., **185** (2014), pp. 560–571. DOI: 10.1016/j.cpc.2013.10.020. arXiv: 1308.3462 [hep-ph].

[131] G. P. Lepage. *A New Algorithm for Adaptive Multidimensional Integration*. J. Comput. Phys., **27** (1978), p. 192. DOI: 10.1016/0021-9991(78)90004-9.

[132] G. P. Lepage. *VEGAS - an adaptive multi-dimensional integration program*. Tech. rep. CLNS-447. Ithaca, NY: Cornell Univ. Lab. Nucl. Stud., 1980.

[133] T. Ohl. *Vegas revisited: Adaptive Monte Carlo integration beyond factorization*. Comput. Phys. Commun., **120** (1999), pp. 13–19. DOI: 10.1016/S0010-4655(99)00209-X. arXiv: hep-ph/9806432.

[134] S. Braß, W. Kilian, and J. Reuter. *Parallel Adaptive Monte Carlo Integration with the Event Generator WHIZARD*. Eur. Phys. J. C, **79**.4 (2019), p. 344. DOI: 10.1140/epjc/s10052-019-6840-2. arXiv: 1811.09711 [hep-ph].

[135] S. Braß. Parallel adaptive Monte Carlo integration and vector-boson scattering at the Large Hadron Collider. PhD thesis. Siegen U., 2019. URN: urn:nbn:de:hbz:467-14658.

[136] T. Ohl. *CIRCE version 1.0: Beam spectra for simulating linear collider physics*. Comput. Phys. Commun., **101** (1997), pp. 269–288. DOI: 10.1016/S0010-4655(96)00167-1. arXiv: hep-ph/9607454.

[137] A. Buckley, J. Ferrando, S. Lloyd, et al. *LHAPDF6: parton density access in the LHC precision era*. Eur. Phys. J. C, **75** (2015), p. 132. DOI: 10.1140/epjc/s10052-015-3318-8. arXiv: 1412.7420 [hep-ph].

[138] W. Kilian, J. Reuter, S. Schmidt, and D. Wiesler. *An Analytic Initial-State Parton Shower*. JHEP, **04** (2012), p. 013. DOI: 10.1007/JHEP04(2012)013. arXiv: 1112.1039 [hep-ph].

[139] S. Braß, W. Kilian, T. Ohl, et al. *Precision Monte Carlo simulations with WHIZARD*. CERN Yellow Reports: Monographs, **3** (2020). Ed. by A. Blondel, J. Gluza, S. Jadach, P. Janot, and T. Riemann, pp. 205–210. DOI: 10.23731/CYRM-2020-003.205.

[140] F. Gaede, T. Behnke, N. Graf, and T. Johnson. *LCIO: A Persistency framework for linear collider simulation studies*. eConf, **C0303241** (2003), TUKT001. arXiv: physics/0306114.

[141] M. Dobbs and J. B. Hansen. *The HepMC C++ Monte Carlo event record for High Energy Physics*. Comput. Phys. Commun., **134** (2001), pp. 41–46. DOI: 10.1016/S0010-4655(00)00189-2.

[142] A. Buckley, P. Ilten, D. Konstantinov, et al. *The HepMC3 event record library for Monte Carlo event generators*. Comput. Phys. Commun., **260** (2021), p. 107310. DOI: 10.1016/j.cpc.2020.107310. arXiv: 1912.08005 [hep-ph].

[143] E. Boos et al. "Generic User Process Interface for Event Generators". *2nd Les Houches Workshop on Physics at TeV Colliders*. 2001. arXiv: hep-ph/0109068.

[144] J. Alwall et al. *A Standard format for Les Houches event files*. Comput. Phys. Commun., **176** (2007), pp. 300–304. DOI: 10.1016/j.cpc.2006.11.010. arXiv: hep-ph/0609017.

[145] J. R. Andersen et al. *Les Houches 2013: Physics at TeV Colliders: Standard Model Working Group Report* (2014). arXiv: 1405.1067 [hep-ph].

[146] W. Kilian and J. Reuter. *The Low-energy structure of little Higgs models*. Phys. Rev. D, **70** (2004), p. 015004. DOI: 10.1103/PhysRevD.70.015004. arXiv: hep-ph/0311095.

[147] W. Kilian, D. Rainwater, and J. Reuter. *Pseudo-axions in little Higgs models*. Phys. Rev. D, **71** (2005), p. 015008. DOI: 10.1103/PhysRevD.71.015008. arXiv: hep-ph/0411213.

[148] W. Kilian, D. Rainwater, and J. Reuter. *Distinguishing little-Higgs product and simple group models at the LHC and ILC*. Phys. Rev. D, **74** (2006). [Erratum: Phys.Rev.D 74, 099905 (2006)], p. 095003. DOI: 10.1103/PhysRevD.74.095003. arXiv: hep-ph/0609119.

[149] J. Reuter and M. Tonini. *Can the 125 GeV Higgs be the Little Higgs?* JHEP, **02** (2013), p. 077. DOI: 10.1007/JHEP02(2013)077. arXiv: 1212.5930 [hep-ph].

[150] J. Reuter, M. Tonini, and M. de Vries. *Littlest Higgs with T-parity: Status and Prospects*. JHEP, **02** (2014), p. 053. DOI: 10.1007/JHEP02(2014)053. arXiv: 1310.2918 [hep-ph].

[151] D. Dercks, G. Moortgat-Pick, J. Reuter, and S. Y. Shim. *The fate of the Littlest Higgs Model with T-parity under 13 TeV LHC Data*. JHEP, **05** (2018), p. 049. DOI: 10.1007/JHEP05(2018)049. arXiv: 1801.06499 [hep-ph].

[152] M. Beyer, W. Kilian, P. Krstonosic, et al. *Determination of New Electroweak Parameters at the ILC - Sensitivity to New Physics*. Eur. Phys. J. C, **48** (2006), pp. 353–388. DOI: 10.1140/epjc/s10052-006-0038-0. arXiv: hep-ph/0604048.

[153] A. Alboteanu, W. Kilian, and J. Reuter. *Resonances and Unitarity in Weak Boson Scattering at the LHC*. JHEP, **11** (2008), p. 010. DOI: 10.1088/1126-6708/2008/11/010. arXiv: 0806.4145 [hep-ph].

[154] W. Kilian, T. Ohl, J. Reuter, and M. Sekulla. *High-Energy Vector Boson Scattering after the Higgs Discovery*. Phys. Rev. D, **91** (2015), p. 096007. DOI: 10.1103/PhysRevD.91.096007. arXiv: 1408.6207 [hep-ph].

[155] W. Kilian, T. Ohl, J. Reuter, and M. Sekulla. *Resonances at the LHC beyond the Higgs boson: The scalar/tensor case*. Phys. Rev. D, **93**.3 (2016), p. 036004. DOI: 10.1103/PhysRevD.93.036004. arXiv: 1511.00022 [hep-ph].

[156] C. Fleper, W. Kilian, J. Reuter, and M. Sekulla. *Scattering of W and Z Bosons at High-Energy Lepton Colliders*. Eur. Phys. J. C, **77**.2 (2017), p. 120. DOI: 10.1140/epjc/s10052-017-4656-5. arXiv: 1607.03030 [hep-ph].

[157] S. Brass, C. Fleper, W. Kilian, J. Reuter, and M. Sekulla. *Transversal Modes and Higgs Bosons in Electroweak Vector-Boson Scattering at the LHC*. Eur. Phys. J. C, **78**.11 (2018), p. 931. DOI: 10.1140/epjc/s10052-018-6398-4. arXiv: 1807.02512 [hep-ph].

[158] N. D. Christensen and C. Duhr. *FeynRules - Feynman rules made easy*. Comput. Phys. Commun., **180** (2009), pp. 1614–1641. DOI: 10.1016/j.cpc.2009.02.018. arXiv: 0806.4194 [hep-ph].

[159] N. D. Christensen, C. Duhr, B. Fuks, J. Reuter, and C. Speckner. *Introducing an interface between WHIZARD and FeynRules*. Eur. Phys. J. C, **72** (2012), p. 1990. DOI: 10.1140/epjc/s10052-012-1990-5. arXiv: 1010.3251 [hep-ph].

[160] F. Staub. *SARAH 4 : A tool for (not only SUSY) model builders*. Comput. Phys. Commun., **185** (2014), pp. 1773–1790. DOI: 10.1016/j.cpc.2014.02.018. arXiv: 1309.7223 [hep-ph].

[161] C. Degrande, C. Duhr, B. Fuks, et al. *UFO - The Universal FeynRules Output*. Comput. Phys. Commun., **183** (2012), pp. 1201–1214. DOI: 10.1016/j.cpc.2012.01.022. arXiv: 1108.2040 [hep-ph].

[162] M. Cacciari, G. P. Salam, and G. Soyez. *FastJet User Manual*. Eur. Phys. J. C, **72** (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].

[163] W. Kilian, J. Reuter, and T. Robens. *NLO Event Generation for Chargino Production at the ILC*. Eur. Phys. J. C, **48** (2006), pp. 389–400. DOI: 10.1140/epjc/s10052-006-0048-y. arXiv: hep-ph/0607127.

[164] T. Robens, J. Kalinowski, K. Rolbiecki, W. Kilian, and J. Reuter. *(N)LO Simulation of Chargino Production and Decay*. Acta Phys. Polon. B, **39** (2008). Ed. by M. Skrzypek and Z. Was, pp. 1705–1714. arXiv: 0803.4161 [hep-ph].

[165] T. Binoth, N. Greiner, A. Guffanti, et al. *Next-to-leading order QCD corrections to pp –> b anti-b b anti-b + X at the LHC: the quark induced case*. Phys. Lett. B, **685** (2010), pp. 293–296. DOI: 10.1016/j.physletb.2010.02.010. arXiv: 0910.4379 [hep-ph].

[166] N. Greiner, A. Guffanti, T. Reiter, and J. Reuter. *NLO QCD corrections to the production of two bottom-antibottom pairs at the LHC*. Phys. Rev. Lett., **107** (2011), p. 102002. DOI: 10.1103/PhysRevLett.107.102002. arXiv: 1105.3624 [hep-ph].

[167] B. Chokoufé Nejad, W. Kilian, J. M. Lindert, et al. *NLO QCD predictions for off-shell $t\bar{t}$ and $t\bar{t}H$ production and decay at a linear collider*. JHEP, **12** (2016), p. 075. DOI: 10.1007/JHEP12(2016)075. arXiv: 1609.03390 [hep-ph].

[168] P. A. Zyla et al. *Review of Particle Physics*. PTEP, **2020**.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.

[169] F. James. *Monte Carlo Theory and Practice*. Rept. Prog. Phys., **43** (1980), p. 1145. DOI: 10.1088/0034-4885/43/9/002.

[170] P. Nason. *A New method for combining NLO QCD with shower Monte Carlo algorithms*. JHEP, **11** (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146.

[171] J. Reuter, F. Bach, B. Chokoufe Nejad, et al. *QCD NLO with Powheg matching and top threshold matching in WHIZARD*. PoS, **RADCOR2015** (2016), p. 088. DOI: 10.22323/1.235.0088. arXiv: 1601.02459 [hep-ph].

[172] J. Reuter, B. Chokoufe, A. Hoang, et al. *Automation of NLO processes and decays and POWHEG matching in WHIZARD*. J. Phys. Conf. Ser., **762**.1 (2016). Ed. by L. Salinas and C. Torres, p. 012059. DOI: 10.1088/1742-6596/762/1/012059. arXiv: 1602.06270 [hep-ph].

[173] C. Bierlich et al. *Robust Independent Validation of Experiment and Theory: Rivet version 3*. SciPost Phys., **8** (2020), p. 026. DOI: 10.21468/SciPostPhys.8.2.026. arXiv: 1912.05451 [hep-ph].

[174] W. Kilian, T. Ohl, J. Reuter, et al. *The WHIZARD public Git repository*. https://gitlab.tp.nt.uni-siegen.de/whizard/public. [Online; accessed 13.06.2021].

[175] T. Gleisberg and S. Hoeche. *Comix, a new matrix element generator*. JHEP, **12** (2008), p. 039. DOI: 10.1088/1126-6708/2008/12/039. arXiv: 0808.3674 [hep-ph].

[176] A. Ballestrero et al. *Precise predictions for same-sign W-boson scattering at the LHC*. Eur. Phys. J. C, **78**.8 (2018), p. 671. DOI: 10.1140/epjc/s10052-018-6136-y. arXiv: 1803.07943 [hep-ph].

[177] J. Alwall, R. Frederix, S. Frixione, et al. *Supplemental run and parameter cards to [33]*. http://amcatnlo.cern.ch/cards_paper.htm. [Online; accessed 13.06.2021].

[178] F. Maltoni, G. Ridolfi, and M. Ubiali. *b-initiated processes at the LHC: a reappraisal*. JHEP, **07** (2012). [Erratum: JHEP 04, 095 (2013)], p. 022. DOI: 10.1007/JHEP04(2013)095. arXiv: 1203.6393 [hep-ph].

[179] R. Frederix, E. Re, and P. Torrielli. *Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO*. JHEP, **09** (2012), p. 130. DOI: 10.1007/JHEP09(2012)130. arXiv: 1207.5391 [hep-ph].

[180] A. Martin, W. Stirling, R. Thorne, and G. Watt. *Parton distributions for the LHC*. Eur. Phys. J. C, **63** (2009), pp. 189–285. DOI: 10.1140/epjc/s10052-009-1072-5. arXiv: 0901.0002 [hep-ph].

[181] M. Cacciari, G. P. Salam, and G. Soyez. *The anti-k_t jet clustering algorithm*. JHEP, **04** (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[182] S. Frixione. *Isolated photons in perturbative QCD*. Phys. Lett. B, **429** (1998), pp. 369–374. DOI: 10.1016/S0370-2693(98)00454-7. arXiv: hep-ph/9801442.

[183] S. Frixione. "Isolated photons without fragmentation contribution". *29th International Conference on High-Energy Physics*. 1998, pp. 790–794. arXiv: hep-ph/9809397.

[184] R. Frederix. *Private Communication* (2019).

[185] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, and G. Heinrich. *NNLO corrections to event shapes in e+ e- annihilation*. JHEP, **12** (2007), p. 094. DOI: 10.1088/1126-6708/2007/12/094. arXiv: 0711.4711 [hep-ph].

[186] S. Weinzierl. *NNLO corrections to 3-jet observables in electron-positron annihilation*. Phys. Rev. Lett., **101** (2008), p. 162001. DOI: 10.1103/PhysRevLett.101.162001. arXiv: 0807.3241 [hep-ph].

[187] N. Baberuxki, C. T. Preuss, D. Reichelt, and S. Schumann. *Resummed predictions for jet-resolution scales in multijet production in e^+e^- annihilation.* JHEP, **04** (2020), p. 112. DOI: 10.1007/JHEP04(2020)112. arXiv: 1912.09396 [hep-ph].

[188] J. Brau, T. Barklow, K. Fujii, et al. *500 GeV ILC Operating Scenarios.* PoS, **ICHEP2016** (2016), p. 062. DOI: 10.22323/1.282.0062.

[189] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. Webber. *New clustering algorithm for multi - jet cross-sections in e^+e^- annihilation.* Phys. Lett. B, **269** (1991), pp. 432–438. DOI: 10.1016/0370-2693(91)90196-W.

[190] A. Buckley, J. Butterworth, L. Lonnblad, et al. *Rivet user manual.* Comput. Phys. Commun., **184** (2013), pp. 2803–2819. DOI: 10.1016/j.cpc.2013.05.021. arXiv: 1003.0694 [hep-ph].

[191] A. Buckley, D. Grellscheid, H. Hoeth, and D. Mallows. *YODA — Yet more Objects for Data Analysis.* https://yoda.hepforge.org/. [Online; accessed 13.06.2021].

[192] A. Buckley, C. Bierlich, C. Gutschow, et al. *Rivet Contrib: Repository with external tools and contributions from and for users of Rivet.* https://gitlab.com/hepcedar/rivetcontrib. [Online; accessed 13.06.2021].

[193] *SLAC Linear Collider Conceptual Design Report* (1980).

[194] M. Vos. *Private Communication* (2021).

[195] S. Kluth, P. A. Movilla Fernandez, S. Bethke, C. Pahl, and P. Pfeifenschneider. *A Measurement of the QCD color factors using event shape distributions at s**(1/2) = 14-GeV to 189-GeV.* Eur. Phys. J. C, **21** (2001), pp. 199–210. DOI: 10.1007/s100520100742. arXiv: hep-ex/0012044.

[196] Y. L. Dokshitzer, G. Marchesini, and B. R. Webber. *Dispersive approach to power behaved contributions in QCD hard processes.* Nucl. Phys. B, **469** (1996), pp. 93–142. DOI: 10.1016/0550-3213(96)00155-1. arXiv: hep-ph/9512336.

[197] Y. L. Dokshitzer and B. R. Webber. *Power corrections to event shape distributions.* Phys. Lett. B, **404** (1997), pp. 321–327. DOI: 10.1016/S0370-2693(97)00573-X. arXiv: hep-ph/9704298.

[198] Y. L. Dokshitzer, A. Lucenti, G. Marchesini, and G. P. Salam. *On the universality of the Milan factor for 1 / Q power corrections to jet shapes.* JHEP, **05** (1998), p. 003. DOI: 10.1088/1126-6708/1998/05/003. arXiv: hep-ph/9802381.

[199] F. Maltoni. *Basics of QCD for the LHC – Lecture II.* https://indico.cern.ch/event/208901/contributions/1501065/. [Online; accessed 13.06.2021]. 2013.

[200] D. P. Barber et al. *Discovery of Three Jet Events and a Test of Quantum Chromodynamics at PETRA Energies*. Phys. Rev. Lett., **43** (1979), p. 830. DOI: 10.1103/PhysRevLett.43.830.

[201] T. Chandramohan and L. Clavelli. *Consequences of 2nd Order QCD for Jet Structure in e^+e^- Annihilation*. Nucl. Phys. B, **184** (1981), pp. 365–380. DOI: 10.1016/0550-3213(81)90224-8.

[202] L. Clavelli and D. Wyler. *Kinematical Bounds on Jet Variables and the Heavy Jet Mass Distribution*. Phys. Lett. B, **103** (1981), pp. 383–387. DOI: 10.1016/0370-2693(81)90248-3.

[203] S. Catani, G. Turnock, and B. Webber. *Jet broadening measures in e^+e^- annihilation*. Phys. Lett. B, **295** (1992), pp. 269–276. DOI: 10.1016/0370-2693(92)91565-Q.

[204] J. D. Bjorken and S. J. Brodsky. *Statistical Model for electron-Positron Annihilation Into Hadrons*. Phys. Rev. D, **1** (1970), pp. 1416–1420. DOI: 10.1103/PhysRevD.1.1416.

[205] 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), pp. 3–42. DOI: 10.1016/0550-3213(93)90271-P.

[206] S. Becker, D. Goetz, C. Reuschle, C. Schwan, and S. Weinzierl. *NLO results for five, six and seven jets in electron-positron annihilation*. Phys. Rev. Lett., **108** (2012), p. 032005. DOI: 10.1103/PhysRevLett.108.032005. arXiv: 1111.1733 [hep-ph].

[207] J. Lindert. *Private Communication* (2020).

[208] B. Chokoufe Nejad, W. Kilian, J. Reuter, and C. Weiss. *Matching NLO QCD Corrections in WHIZARD with the POWHEG scheme*. PoS, **EPS-HEP2015** (2015), p. 317. DOI: 10.22323/1.234.0317. arXiv: 1510.02739 [hep-ph].

[209] M. Czakon. *A novel subtraction scheme for double-real radiation at NNLO*. Phys. Lett. B, **693** (2010), pp. 259–268. DOI: 10.1016/j.physletb.2010.08.036. arXiv: 1005.0274 [hep-ph].

[210] M. Czakon and D. Heymes. *Four-dimensional formulation of the sector-improved residue subtraction scheme*. Nucl. Phys. B, **890** (2014), pp. 152–227. DOI: 10.1016/j.nuclphysb.2014.11.006. arXiv: 1408.2500 [hep-ph].

[211] H. A. Chawdhry, M. Czakon, A. Mitov, and R. Poncelet. *NNLO QCD corrections to diphoton production with an additional jet at the LHC* (2021). arXiv: 2105.06940 [hep-ph].

[212] S. Braß, P. Bredt, W. Kilian, et al. "Complete Automation of NLO SM Processes in WHIZARD". in preparation.

[213] S. Catani, M. H. Seymour, and Z. Trocsanyi. *Regularization scheme independence and unitarity in QCD cross-sections*. Phys. Rev. D, **55** (1997), pp. 6819–6829. DOI: 10.1103/PhysRevD.55.6819. arXiv: hep-ph/9610553.

[214] E. Bothmann, S. Hoeche, F. Krauss, et al. *SHERPA 2.2.10 Manual*. https://sherpa.hepforge.org/doc/SHERPA-MC-2.2.10.html. [Online; accessed 13.06.2021].

Acknowledgements

During the last few years I worked towards and on this thesis, I enjoyed collaborations and scientific discussions with many people, with overall support by many more. I am grateful for every one of them and would like to use this opportunity to thank some of them explicitly.

First of all, I would like to thank my supervisor Jürgen Reuter for the possibility to be a PhD student as part of his group. He gave invaluable scientific advice whenever I needed it and fruitful discussions with him were plentiful. I am very thankful for his overwhelming support during the process of writing this thesis.

I would like to thank Sven-Olaf Moch for being my second supervisor and for answering questions about the administrative side of performing my PhD.

I would like to thank Caren Hagner for chairing my thesis committee, as well as Markus Diehl and Peter Schleper for joining it.

Furthermore, I would like to give very special thanks to Christian Weiss and Bijan Chokoufe Nejad for paving the way for this thesis by their previous work and for giving enlightening insights into the complexity of the WHIZARD source code. Well after his departure from academia, Christian was always eager to answer remaining technical questions. Bijan introduced me to the quirks of Git and GitLab, allowing me to continue the support for the WHIZARD group after his departure.

I am very grateful for my WHIZARD and Deutsches Elektronen-Synchrotron (DESY) colleagues Simon Braß, Pia Bredt, and Pascal Stienemeier. Together, we formed an effective team to squat many code bugs, had uncountable constructive discussions, and significantly shaped the form of WHIZARD. Pascal helped me in finding a few especially elusive bugs in the computation of total cross sections and supported me in the validation effort. Both Pascal and Pia were invaluable in discussions about fNLO event generation, finding bugs therein, and validating various aspects of it. Without Simon's effort to parallelize WHIZARD and give continuous support, however, this thesis would not have been possible within the given time frame. In addition, he gave invaluable support in redefining internal mappings in WHIZARD. I would like to thank all three for of proofreading this thesis.

Additionally, I owe many thanks to Wolfgang Kilian and Thorsten Ohl for valuable scientific and technical input regarding questions about physics and WHIZARD.

I would like to express my thanks to the complete WHIZARD team for many

productive discussions and meetings, as well as enjoyable evenings in Hamburg, Siegen, and Würzburg.

I would like to thank Jonas Lindert for his support with OpenLoops and providing the library for $e^+e^- \to jjjjjj$.

I would like to thank Rikkert Frederix for answering many questions regarding MG5_aMC@NLO and FKS subtraction.

I would like to thank DESY for the financial support during my PhD studies in terms of salary and making it possible to visit several summer schools. I am very grateful for the opportunity given to me by the DESY IT group to use the Maxwell high-performance computing cluster. Without it, many computations for this thesis would not have been possible in time.

I am grateful for all the administrative support given by the secretaries of the DESY theory group, Cristina Guerrero and Julia Herrmann.

I would like to thank Anne Ernst and Yannick Linke for being a great team during my time in the SFB 676 Student Council.

My office colleagues – Anne Ernst, Ivan Sobolev, Troy Figiel, and Daniel Meuser – sparked many discussions about science and general topics, as well as many laughs, for which I am very grateful.

I am thankful for all the great lunch times with the obligatory coffee breaks which have been enriched by discussions with Henning Bahl, Tim Stefaniak, and Jonas Wittbrodt. I would like to thank Jonas for proofreading this thesis.

I am very grateful for all the wonderful evenings with my friends of the DESY theory board game evenings, which have been great fun and a perfect distraction.

I owe very special thanks to Rachel Fulda who endured my long journey of working towards and finishing this thesis. Her support and encouragement brought me through many stressful times and without her I would certainly have had a much harder time finishing my PhD studies.

Last, but not least, I would like to thank my family, especially my parents Birgit Rothe and Jan Hauschildt, for the heartfelt support throughout my studies prior and during my time as a PhD student. Without them, I would not have been able to make it this far.

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt zu haben.

Hamburg, den 14.06.2021

Vincent Rothe