Improving the Efficiency and Accuracy in Monte-Carlo Event Generation

A thesis submitted for the degree of

Doctor of Philosophy

by

Christian Tobias Preuss
BSc, MSc

School of Physics and Astronomy
Monash University
Australia

August 2021

Copyright Notices

Notice 1

Notice 2

To my family,

the colour in my life.

Contents

Abstract

While statistical uncertainties at the Large Hadron Collider (LHC) continue to decrease, uncertainties pertaining to theoretical modelling become more and more dominant, demanding higher precision in event-generation tools. At the same time, it has become evident that the computing power available to LHC experiments is unlikely to meet the demand during the high-luminosity phase of the LHC. For the discrimination of new-physics signals from (known) Standard-Model phenomena, detailed knowledge of unwanted background processes is necessary with high statistical significance. The precision and efficiency constraints of current event generators therefore hinder the full realisation of the physics potential of the LHC as well as future colliders.

Parton showers are among the most widely used theoretical tools in collider physics and core part of any multi-purpose Monte Carlo event generator. Despite their vital importance in the modelling of so-called logarithmically enhanced radiation, these algorithms employ a series of approximations yielding only limited accuracy in the description of highly-complex collider events. While the situation may be improved upon by the systematic combination with accurate (and often precise) perturbative calculations, via procedures known as matching and multi-jet merging schemes, typically, the complexity of such algorithms render these inefficient for events with many particles. Moreover, the combination of parton-shower simulations with precision calculations is so far limited to the next-to-leading order in a perturbative series of the strong coupling constant with first approaches extending it to the next-to-next-to-leading order for a few processes.

This thesis focusses on a so-far little explored parton-shower framework called sector showers, which is designed for the efficient combination with higher-order perturbative corrections. Within the sector-shower approach to parton showers, the branching phase space is divided into non-overlapping regions, each of which corresponds to a certain parton-shower branching. This drastically reduces the usually factorially growing number of so-called parton-shower histories to only a single history in best cases.

A new and full-fledged implementation of sector showers in the VINCIA antenna shower in PYTHIA 8.3 is presented alongside a dedicated multi-jet merging scheme. It is demonstrated that the maximally bijective nature of sector showers directly translates into a drastically improved scaling of the event-generation time and memory footprint with the particle multiplicity. First steps towards and efficient algorithm to match the new shower implementation with fixed-order calculations are presented. To this end, it will be demonstrated how the use of analytic forms of so-called one-loop amplitudes can significantly accelerate state-of-the-art precision calculations. Facilitated by the use of sector showers, a proof-of-concept implementation of a novel method to match parton showers to second-order calculations will be outlined. As a proof of the usability of the revised VINCIA antenna shower in real-life setups, a study of Higgs production in the phenomenologically interesting vector boson fusion channel is presented.

Publication List

This thesis is based on the following works:

I Helen Brooks, Christian T Preuss, and Peter Skands, *Sector Showers for Hadron Collisions*, JHEP 07 (2020) 032, doi:10.1007/JHEP07(2020)032, [arXiv:2003.00702]

II Helen Brooks, Christian T Preuss, *Efficient multi-jet merging with the Vincia sector shower*, CPC 264 (2021) 107985, doi:10.1016/j.cpc.2021.107985, [arXiv:2008.09468]

III Stefan Höche, Stephen Mrenna, Shay Payne, Christian T Preuss, and Peter Skands, *A Study of QCD Radiation in VBF Higgs Production with* VINCIA *and* PYTHIA, submitted to SciPost Physics, [arXiv:2106.10987]

IV John M Campbell, Stefan Höche, and Christian T Preuss, *Accelerating LHC phenomenology with analytic one-loop amplitudes: A C++ interface to MCFM*, submitted to EPJC, [arXiv:2107.04472]

V John M Campbell, Stefan Höche, Hai Tao Li, Christian T Preuss, and Peter Skands, *Towards NNLO+PS Matching with Sector Showers*, submitted to PLB, [arXiv:2108.07133]

Additionally, the following work has been published during the candidature, but is not included in this thesis:

VI Nick Baberuxki, Christian T Preuss, Daniel Reichelt, Steffen Schumann, *Resummed predictions for jet-resolution scales in multijet production in e^+e^- annihilation*, JHEP 04 (2020) 112, doi:10.1007/JHEP04(2020)112, [arXiv:1912.09396]

Declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes 2 original papers published in peer reviewed journals and 3 submitted publications. The core theme of the thesis is the improvement of the efficiency and precision of particle-level event simulations in Monte Carlo event generators. The ideas, development, and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working within the School of Physics and Astronomy under the supervision of Peter Skands.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

My contribution to each of the included works is stated below.

Publication title *Sector Showers for Hadron Collisions*
Status Published, Thesis Chapter 4
Student contribution (*75%*)
Developed the initial-state sector criteria. Wrote the code implementation and validated it. Performed the phenomenological studies of initial-state and final-state sector showers. Wrote the paper draft.
Collaborator contributions
Helen Brooks (*10%*) Helped with the implementation of resonance-final sector showers and performed the phenomenological study with them. Contributed to discussions and the write-up.
Peter Skands (*15%*) Developed the idea and implemented the antenna functions. Responsible for overall project management. Contributed to discussions and the write-up.

Publication title *Efficient multi-jet merging with the Vincia sector shower*
Status Published, Thesis Chapter 5
Student contribution (*75%*)
Contributed to the code implementation. Validated the merging implementation and performed the speed and memory benchmarking. Performed the phenomenological study. Wrote the paper draft.
Collaborator contribution
Helen Brooks (*25%*) Developed the underlying code structure for the merging implementation. Contributed to discussions and proof-read the write-up.

Publication title *Accelerating LHC Phenomenology with analytic one-loop amplitudes: A C++ Interface to* MCFM
Status Submitted, Thesis Chapter 6
Student contribution (*50%*)
Initiated the development and lead the collaboration. Contributed to the code implementation and validation. Performed the benchmarking. Contributed to the write-up.
Collaborator contributions
John M Campbell (*25%*) Developed and validated the code implementation. Performed cross-checks with MCFM. Contributed to discussions and the write-up.
Stefan Höche (*25%*) Developed and validated the code implementation. Performed the benchmarking. Contributed to discussions and the write-up.

Publication title *Towards NNLO+PS Matching with Sector Showers*
Status Submitted, Thesis Chapter 7
Student contribution (*60%*)
Developed the NNLO matching formalism. Wrote the code implementation of the antenna subtraction method and contributed to the implementation of two-to-four showers. Implemented matrix-element corrections and the COMIX interface. Validated the code implementation. Initiated and lead the collaboration; organised regular meetings. Wrote the paper draft.
Collaborator contributions
John M Campbell (*8%*) Helped with the development of the subtraction method and analytical integration of antenna functions. Implemented analytic matrix elements. Contributed to discussions and the write-up.
Stefan Höche (*10%*) Helped with the development of the matching method and analytical integration of antenna functions. Implemented the COMIX interface. Contributed to discussions and the write-up.
Hai Tao Li (*10%*) Originally developed the two-to-four antenna shower framework and contributed to its re-implementation in VINCIA. Validated the code implementation. Contributed to discussions and proof-read the write-up.
Peter Skands (*12%*) Originally developed the two-to-four antenna shower framework and the NLO matrix-element correction method. Implemented the basic matrix-element correction framework. Contributed to discussions and the write-up.

Publication title *A Study of QCD Radiation in VBF Higgs Production with* VINCIA *and* PYTHIA
Status Submitted, Thesis Chapter 8
Student contribution (*70%*)
Extended the merging implementation for VBF processes, wrote the POWHEG hooks for VINCIA. Lead the phenomenological study. Generated the POWHEGBOX event samples and performed the leading-order event generation using SHERPA event samples and PYTHIA-internal events. Performed the study of non-perturbative effects. Produced all plots and wrote the paper draft.
Collaborator contributions
Stefan Höche (*6%*) Generated the SHERPA event samples. Contributed to discussions and the write-up.
Stephen Mrenna (*6%*) Implemented the parallelisation for VINCIA. Performed the tree-level merged event generation with SHERPA event samples.
*Shay Payne** (*3%*) Helped with the NLO-matched event generation with POWHEG event samples. Contributed to the write-up.
Peter Skands (*15%*) Developed the original idea for the study. Defined and suggested new observables. Contributed to discussions and the write-up.

* marks co-authors who were Monash students while part of the work was undertaken.

I have not renumbered sections of submitted or published papers.

Student Name: Christian Tobias Preuss **Date:** August 26, 2021

I hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors

Main Supervisor Name: Peter Skands **Date:** August 26, 2021

Acknowledgements

What a journey this has been – both literally and figuratively. The past years have certainly been very different to what I had imagined, but it has nevertheless been a tremendous and unforgettable experience. This thesis would not be what it became without the help and work of so many people who accompanied this journey.

First and foremost, I want to express my gratitude to my supervisor Peter Skands. It has been an exceptional pleasure to work with him and I am glad to have had him as my supervisor, mentor, and friend. Without Peter's support and advice from the very first second, not only on physics but life in general, I would not have been able to complete my degree the way I did. I have benefited from Peter's vast knowledge and physics intuition. I would like to thank him for giving me the freedom to bring in my own ideas, for encouraging me to follow my intuition, for allowing me to attend the CTEQ school in Pittsburgh and to spend time at Fermilab, for encouraging me to collaborate with other academics, for staying excited when I was critical and for being critical when I was excited.

I would also like to thank my associate supervisor German Valencia for his continuous encouragement to work on BSM modelling. The chats we had were always stimulating and I have learned many new things along the way. The atmosphere in the HEP group at Monash has made my candidature especially pleasant. The regular group meetings, coffee (and cake) sessions, and student get-togethers have always been terrific, both in person and online. I would like to thank everyone in the group for making it possible for me to regularly join these while I was overseas. I want to thank my fellow postgrads and especially Cody Duncan for the great atmosphere in the cohort. My work would not have been possible without the administrative help by Jean Pettigrew and Karen Lee and the financial support by Monash University through the Monash Graduate Scholarship, the Monash International Postgraduate Research Scholarship, and the J.L. William Scholarship.

My first year at Monash would not have been the same without Helen Brooks, whose office door was always open for me and who patiently answered all my questions on showers, VINCIA, PYTHIA, and coding. I am thankful to her for teaching me the art of merging and for her help with the implementation of the *MESS*.

Steffen Schumann deserves special mention here, as it is due to him that I have chosen to work on phenomenology. Without his encouragement to move away for a PhD I would not have chosen to go to Monash, a decision I am exceptionally happy to have made. I would also like to thank Daniel Reichelt for many discussions on colour correlations and resummation and for chauffeuring me through the suburbs of Chicago.

I am fortunate to have had the opportunity to work with John Campbell and Stefan Höche. Their interest in my work, the countless discussions we had online, and their immense expertise and experience have shaped my understanding of fixed-order calculations,

subtraction schemes, and matching. It truly felt like having two additional supervisors and I would like to thank both of them for their support. I am also glad to have worked with Hai Tao Li, whose help with the implementation and debugging of the two-to-four branching framework was indispensable.

During my candidature, I had the chance to join the PYTHIA collaboration and I would like to thank everyone in the collaboration not only for accepting me as a new member, but also for stimulating meetings and PYTHIA weeks. Especially, I want to thank Philip Ilten for thorough reviews of my code, Stephen Mrenna for help with parallelisation, and Stefan Prestel for interesting and helpful chats on matching and merging.

My friends, both in Europe and Australia, should be mentioned here. Especially, I would like to thank Benedikt Ringbeck for his strong and long-lasting friendship and for his comments on the introduction of this thesis. Without Jake Bartels I wouldn't have settled in Melbourne as quickly and comfortably as I did and he was the best flatmate and friend I could have wished for.

Finally, I cannot express how grateful I am to have a supporting and loving family. My sister Dorothee Preuss who made my childhood and youth as exceptional as it it could be and with whom I can always talk about everything I have on my heart. My *Mama*, Angela Preuß, who fostered my independence, my critical thinking, and encouraged me to philosophise about politics, nature, and religion – I will never forget the myriad late-night discussions with her about *Gott und die Welt*. My *Opa*, Rudolf Preuß, who taught me the value of education and who never hesitated to support my education. My *Schwiegereltern*, Birgit and Uwe Lemties, who always made me feel welcome to the family and who fueled me with *Nervennahrung* during my last year of candidature. Most importantly, I owe my partner Luisa Lemties a debt of gratitude. The past years have certainly not always been easy, living a world away from each other for most of the time and having to stand through the world locking down while being separated. Her patience with me, her unconditional support in every situation, and her laid-back character have guided me through life over the past years. It is beyond words how fortunate I am to share my life with Luisa and how incredibly happy I am about our beautiful baby daughter Leah. Although I have not had the time I wished I had for both of them during the writing of this thesis, calling them my family is one of the best things in my life.

Put something silly in the world
That ain't been there before.
Shel Silverstein *Put Something In*

1

Introduction

Curiosity is one of the most natural drives of human life. When receiving a present, the first thought that crosses one's mind is »what's in there?«. In a way, particle physicists are just like children on their birthdays, seeing a huge table full of nicely-wrapped presents thinking of nothing else but this very question. Unarguably, there is a difference: the gift is nature itself and it neither comes served on a table nor neatly wrapped in shiny paper. The question, however, stays the same: »what's in there?«. Let us take it as a guide into the fascinating world of *Elementary Particle Physics*.

When zooming in on what *matter* is comprised of, *cf.* fig. 1.1, we will first see *molecules* emerge, which in turn are bound states of multiple *atoms*. The nature of atoms was for a long time subject of philosophical discussion. While in ancient Greek natural philosophy, the word *atomos* already reflected the understanding of a "granular" nature of matter, being built from "uncuttable" objects [1], it was later observed by English chemist John Dalton that different chemical compounds containing the same element come in constant mass ratios, formulated in terms of the *law of multiple proportions*. This confirmed that each element is indeed made up of a single type of atoms with a specific mass. The theory of "uncuttable" atoms then underwent a rapid series of changes around the turn of the $19^{\underline{\text{th}}}$ and $20^{\underline{\text{th}}}$ century. It was Sir J.J. Thomson who discovered the existence of "corpuscles", small charged particles which are lighter[1] than atoms [2], today known as *electrons*. As he discovered these new particles in cathode rays, he came to the conclusion that they must be *subatomic*, meaning they form parts of the atoms in the cathode material – the end of the idea of the indivisable atom. The atomic model he described is often referred to as the "plum pudding" model: negatively charged electrons ("plums") sit uniformly distributed inside the positively charged atom "pudding". This perspective was disproved by the Rutherford scattering experiments[2], in which alpha particles were shot on gold foil [3, 4]. Different to what one would expect from Thomson's model, a proportion of particles were deflected to large angles. This led Rutherford to the proposal of an atomic *nucleus*[3], in which all of the positive charge is concentrated, surrounded by a sphere of electrons [5]. With the positive charges centred in the atom, positively charged alpha particles (themselves being helium-4 nuclei) are deflected by the nucleus' electric field. As quantum theory emerged [6, 7] and it was observed that atoms emit light in discreet frequencies only [8–11], it became clear that such a "planetary" model of the atom had

[1] Thomson himself reported them to be *smaller* than atoms, which from today's perspective of point-like electrons seems ill-defined.

[2] The experiments were actually conducted by Hans Geiger and Ernest Marsden.

[3] The physical term nucleus is borrowed from the biological term for the cell nucleus.

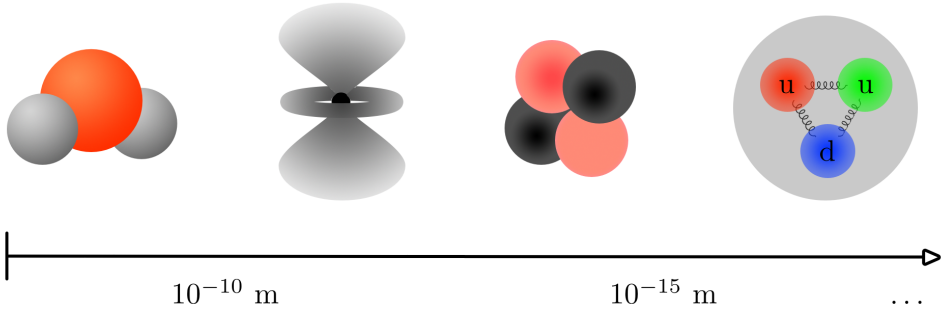


Figure 1.1: *From left to right: illustration of a water molecule, a hydrogen atom with electron orbitals, a helium-4 nucleus (α particle), a proton consisting of two up and one down quark held together by gluons (curly lines). Length scales not to scale.*

critical deficiencies. If electrons really orbit the atomic nucleus, they necessarily must emit *bremsstrahlung* as a consequence of them being accelerated charges. This not only means that they emit light with a continuous spectrum, but more direly that eventually they will lose all their energy and fall right into the nucleus, making atoms unstable. This issue is circumvented in the Bohr model, named after Danish physicist Niels Bohr, which postulates that electron orbits are *quantised* [12], meaning that only those "stationary" orbits with radii of integer multiples of the Planck constant $\hbar$ [4] are allowed in order for the atoms to remain stable. The criterion of discrete circular orbits was later loosened to ellipses by Arnold Sommerfeld [13]. With the formulation of Quantum Mechanics [14–17], the Bohr model got a new spin: electrons do not travel on orbits, but instead surround the nucleus as a "cloud" called *orbital*, corresponding to the probability density of the electron's position. In 1932, Carl Anderson discovered a "positive electron" in cosmic ray tracks in a cloud chamber [18], theorised before by Paul Dirac as an implication of his relativistic description of the electron [19]. Today, we call this *antiparticle* of the electron the *positron* [20] and know that to each matter particle there exists an *antimatter* particle.

After the discovery of the neutron by James Chadwick in 1932 [21, 22], models of the nucleus were developed containing both, positively charged *protons* and electrically neutral *neutrons* [23–26]. Because protons repel each other by the electric force, they must be bound together in the nucleus by another stronger force, hence dubbed the *strong force*. It was then observed that the strong force had a *symmetry*: it treats both protons and neutrons in the same way (different to the electromagnetic force, which only acts on protons). Moreover, Murray Gell-Mann and George Zweig independently proposed the *quark model* [27, 28], describing protons and neutrons not as fundamental particles but composite objects of three new particles called *quarks*. The term *quark* was coined by Murray Gell-Mann, who took inspiration from the Irish novelist James Joyce's illegible work *Finnegan's Wake*, in which the following lines can be found [29]:

> *Three quarks for Muster Mark!*
> *Sure he hasn't got much of a bark*
> *And sure any he has it's all beside the mark.*

The quark model may be seen as a generalisation of the proton-neutron symmetry mentioned above, but assumes a symmetry between three quark "*flavours*" up (u), down (d),

[4] The Planck constant $h = 6.62607015 \times 10^{-34}$ m^{2}kgs^{-1}, named after German physicist Max Planck who introduced it in [6, 7], is the fundamental constant relating the frequency of a photon to its energy, $E_\gamma = h\nu_\gamma$. The quoted constant $\hbar = h/2\pi$ denotes the *reduced* Planck constant.

and strange (s), *i.e.*, that there is no way to distinguish the three[5]. The existence of this symmetry has a striking feature: it provides a rigorous and systematic way to derive the composition and properties of other *hadrons*, *i.e.*, particles comprised of a combination of quarks and/or antiquarks. Today, we know that the up, down, and strange quarks are only the three lightest of a total of six quark flavours, with an additional three *heavy quarks* called charm (c), bottom (b), and top (t). The latter is almost 200 times heavier than a hydrogen atom and was discovered in 1995 at Fermilab's Tevatron [30, 31].

Some atomic nuclei are unstable and decay via the emission of alpha particles (α decay), an electron and an electron-antineutrino (β decay), or high-energy photons (γ decay). We have seen above how the study of α particles shaped the understanding of atoms, eventually leading us to the strong force which keeps atomic nuclei together. Similarly, studies of β decay shaped our understanding of the *weak force*. In order to save energy-momentum conservation in β decay, Wolfgang Pauli hypothesised the existence of a new particle in an open letter to Lise Meitner [32]. He proposed that the electron emitted in the decay is accompanied by an extremely light particle – the neutrino[6] – which could take the necessary recoil in the reaction. This idea was taken up by Enrico Fermi, who formulated a first theory of this type of interaction, represented by a four-point vertex between the proton, the neutron, the electron, and the anti-neutrino in the decay [33]. The existence of the so-far hypothetical neutrino was only shown about twenty years later by Frederick Reines and Clyde Cowan [34]. Today, we know that *leptons*, just as quarks, come in six flavours; there are three charged leptons: the electron (e^-), the muon (μ^-), and the tau-lepton (τ^-); and three neutral leptons: the electron-neutrino (ν_e), the muon-neutrino (ν_μ), and the tau-neutrino (ν_τ).

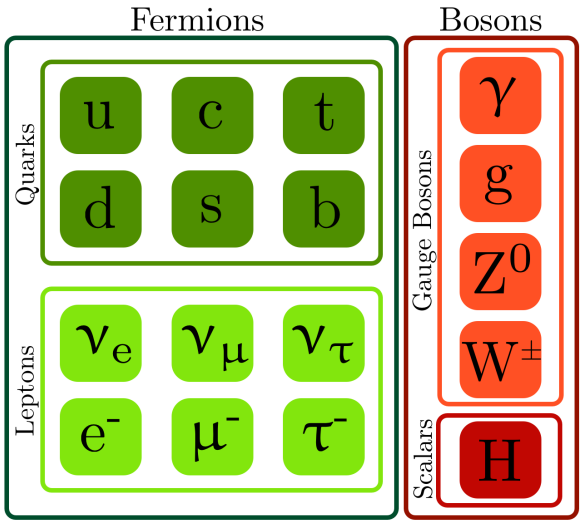


Figure 1.2: *The particle content of the Standard Model.*

In modern particle physics, the understanding of subatomic particles is condensed in the *Standard Model of Particle Physics*. Within the Standard Model, matter particles are described in terms of quasi-localised[7] point-like excitations of a *quantum field*. The particle content of the Standard Model is depicted in fig. 1.2. Interactions between particles are formulated in terms of *local gauge symmetries*, imposing that observable phenomena be *invariant* (*symmetric*) under a *local redefinition* (*gauge*) of the unobservable quantum fields. Simply spoken, any measurement must not depend on the specific choice of unmeasurable quantities. The Standard Model contains three such local gauge symmetries with the catch that each of them gives rise to *force carrier* particles in the form of *gauge bosons*. Within the canonical interpretation of quantum field theory, gauge bosons are thought of as *mediating* the interaction between matter particles. As an example, photons are considered to mediate the electromagnetic force, formulated as Quantum Electrodynamics (QED) in the Standard Model. The aforementioned weak and strong

[5] This symmetry holds only if the masses of the three quarks are identical. As this is not the case, but the masses are of similar magnitude, this symmetry holds only approximately.

[6] Interestingly, Pauli first called it neutron.

[7] The *uncertainty relation* formulated by Werner Heisenberg in [35] forbids the simultaneous precise measurement of the position and momentum of any object.

force are mediated by the $W^{\pm}$ and Z^0 bosons and eight gluons, respectively. As such, it is the W^- boson that mediates the β^- decay reaction $n \to p + (W^- \to e^- + \bar{\nu}_e^-)$ and the gluons are responsible for "gluing" the quarks in hadrons together. The latter were observed in electron-positron collisions at the DESY Positron-Electron Tandem Ring Accelerator (PETRA) in 1979 [36–38], whereas the existence of the weak bosons $W^{\pm}$ and Z^0 was observed in 1983 in measurements of proton-antiproton collisions at the CERN Super Proton Synchrotron (SPS) [39,40]. Probably the most famous particle is the Higgs, which contrary to the gluon, Z^0, and $W^{\pm}$ is not a gauge boson, *i.e.*, does not mediate a force. Instead, fundamental particles (except for neutrinos which are assumed massles in the Standard Model) obtain their masses by interacting with the Higgs field. After years of searching, the Higgs was finally discovered at the CERN Large Hadron Collider (LHC) in 2012 [41,42]. It concluded the search for the Standard-Model particles and added the last piece to the theory.

Although the Standard Model has impressively proven its own success, it cannot be the whole story. To name only a few shortcomings, it is now known that neutrinos must have a non-zero mass, while they are predicted exactly massless in the Standard Model; the ordinary matter makes up only a few percent of the matter in the universe; the Standard Model describes only three of the four forces we deem fundamental today, but leaves out gravity. Hence, we are not merely asking the question »what's in there?« but rather the humble questions »where does everything come from and how does the universe work?«. To address these, further research for *new physics* and *physics beyond the Standard Model* is necessary.

The history of (subatomic) particle physics is also a history of *particle accelerators* and *particle collisions*, with simple cathode rays at the beginning of the story and large-scale colliders, such as the LHC shown in fig. 1.3, at the end. So why do we need to accelerate and collide particles for our understanding of the universe? As smaller and smaller length scales are probed, increasingly small wave lengths are needed, corresponding to increasingly large energies. This is the reason why particle physics today is often also referred to as *high-energy physics*. Colliders are large linear or circular tubes, in which (typically two[8]) particle beams are accelerated to velocities close to the speed of light and then collided at certain interaction points. The products of reactions between the beam particles are then measured by detectors surrounding the beam pipe at these interaction points.

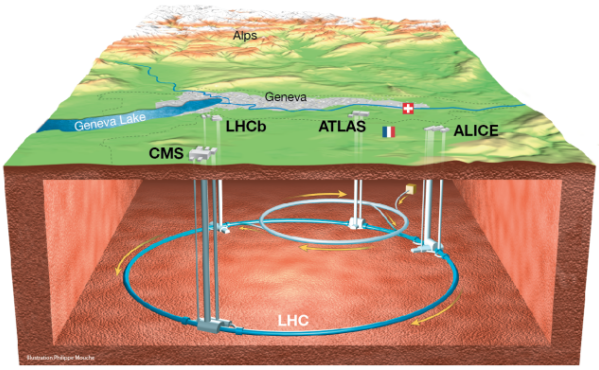


Figure 1.3: *Overview of the LHC and its experiments. Taken from* [43].

On first sight there seems to be a wide gap between the mathematically complex Standard Model and the reconstructed tracks of particles in detectors of high energy physics experiments. This is where Monte Carlo event generators come into play. Event generators are multi-purpose tools to simulate full collider *events*, *i.e.*, everything from the collision of two beams to the creation of all the observable final-state particles, just in the same way as one would "see" it in a detector. The "Monte Carlo" part enters the name of event generators, as these *probabilistically* generate particle configurations using

[8] In *fixed-target* experiments, only one beam is accelerated and shot on a stationary target. This is energetically unfavourable, so that typically two accelerated beams are used.

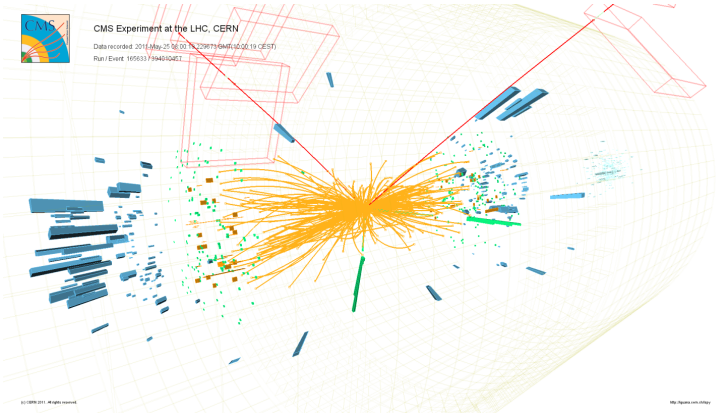


Figure 1.4: *Actual event display of the production of two* Z^0 *bosons in a proton-proton collision measured by the CMS experiment. Taken from* [44].

a technique known as *Monte Carlo sampling*. It provides a simple means to calculate arbitrarily complicated integrals (or generalised volumes) by a "hit-and-miss" approach, meaning that the volume of a certain object can be calculated by randomly throwing points in an encompassing region and counting how many times these hit the object. To provide such a detailed description, a particle collision is roughly divided into three stages: the *hard scattering* in which two beam particles collide to create a small number of highly-energetic particles; the *parton shower* during which the particles of the hard scattering lose energy via the radiation of softer particles kown as *bremsstrahlung*; the *hadronisation* phase in which the large number of low-energy particles after the shower are combined into observable hadrons. This factorisation of the event corresponds to a "time ordering" in terms of the energy, *i.e.*, high energies (*hard* phenomena) appear early in the simulation, while low energies (*soft* phenomena) appear late. While the hard reaction can be calculated up to a certain order in a perturbative series dictated by the underlying Quantum Field Theory, parton showers are based on the radiation pattern of fundamental particles, in turn derived from Quantum Field Theory. This allows to describe the entire cascade from the few particles in the hard scattering to very many soft particles, which can be combined into visible hadrons. Different to the previous two parts, the last part of the simulation, the *hadronisation*, cannot be derived from perturbative Quantum Field Theory. It is, instead, based on judicious assumptions about the behaviour of quarks and gluons at low energies. As such, there exist models of quite different nature to describe these phenomena. Through their *exclusive* description of collider signals, event generators have become an indispensable tool for particle physics, without which neither data analysis nor comparison of experimental and theoretical predictions would be possible beyond a qualitative level.

To facilitate the search for new physics phenomena, a precise knowledge of the standard-model *background* is necessary. It has thus become clear that higher-precision event generators are needed to meet the increasing experimental precision of current colliders. At the same time, the intense use of these tools implicates severe computing challenges for future physics programs. Projections show that the computing power of the LHC experiments falls short of the need for high-statistics simulations during the high-luminosity phase of the LHC [45–48]. The precision and efficiency constraints of current event generators therefore hinder the full exploitation of the physics potential of the LHC as well as future colliders. It is thus evident that new and refined models are needed to meet the demand for higher-precision as well as higher-efficiency event generation tools. This thesis aims at addressing parts of these limitations. It is centred around the development of a novel parton-shower algorithm dubbed *sector showers* and the systematic and efficient correction of their accuracy and precision.

Outline of the thesis

After a brief recapitulation of Quantum Field Theory and the Standard Model of Particle Physics in chapter 2 and a review of Particle Physics Phenomenology and Monte Carlo event generators in chapter 3, the implementation of sector showers in the VINCIA antenna shower in PYTHIA 8.3 will be described in detail in chapter 4. Chapter 5 presents a multi-jet merging approach utilising sector showers in order to drastically mitigate the computational and memory overhead typically inherent to such techniques. In chapter 6, a new generic interface to the extensive library of analytic one-loop matrix elements in the MCFM parton-level event generator is presented and the timing improvements compared to existing automated one-loop providers demonstrated. The precision frontier is addressed in chapter 7, where, for the first time, a proof-of-concept implementation of a fully-differential next-to-next-to-leading order plus parton-shower matching scheme is presented, using the sector shower implementation in VINCIA and the MCFM one-loop matrix element interface. In chapter 8, VINCIA's sector showers are tested in a realistic phenomenological study of Higgs boson production via vector boson fusion, employing the multi-jet merging scheme of chapter 5 and comparing it to state-of-the-art next-to-leading order matched predictions. Chapter 9 summarises the main points of this thesis, puts it into the bigger perspective, and gives an outlook on future developments.

2

QFT and the Standard Model

In this chapter, we will describe the mathematical foundations of the *Standard Model of Particle Physics*, herein often referred to simply as the *Standard Model* (SM). The SM is constructed as a *local gauge theory* within Quantum Field Theory (QFT). Since the latter is a mature and complex field in itself, only a concise overview can be given here. The interested reader is referred to text books on Particle Physics [49–52] and QFT [53–57], covering the subject at various levels. Nevertheless, a brief review of Quantum Field Theory and local gauge theories will be given in section 2.1. The SM as well as its quantum-field theoretic description will be discussed in section 2.2.

This chapter shall serve two purposes: firstly, it shall give a concise overview over the mathematical foundations of particle physics; secondly, it shall provide the starting point for the discussion of Monte Carlo event generators in chapter 3. Some of the material covered in the latter chapter could therefore equally well be discussed here. This is not done, so that a broader perspective can be retained within this chapter, while a more rigorous discussion can be postponed to chapter 3, although some of the mathematical discussions here may break this rule of thumb.

2.1 QFT, Lie Groups, and Yang-Mills Theory in a Nutshell

In particle physics, we deal with tiny, sub-atomic particles at very high energies. To obtain a faithful theoretical description of it, we therefore have to combine two of the major breakthroughs of early-20th century physics into a single theory: Quantum Mechanics (QM) and Special Relativity (SR). Furthermore, as we know that particles can have wave-like properties and vice versa (a fact known as the particle-wave duality), we think of particles and waves merely as two different realisations of a single entity, a field. A field is a quantity which extends over all of space and time, hence possessing infinitely many degrees of freedom. It can have different modes of excitation: a well-localised excitation with a sharp peak or an extended modulation of its amplitude without a clear definition of a peak structure. We will think of the former as a *particle* and the latter as a *wave*.

The dynamics of fields, *i.e.*, the variation of their amplitude in space and time, is understood in terms of an *action*,

$$S[\phi] = \int_{t_1}^{t_0} \mathrm{d}t \int \mathrm{d}^3x \, \mathcal{L}(\phi(x), \partial_\mu \phi(x)) \,, \tag{2.1}$$

which is a *functional*, mapping the *Lagrangian density* $\mathcal{L}$ (hereinafter simply referred to as the *Lagrangian*) to a scalar S. The action integral may be understood as a quantity which takes all possible configurations of a given system into account when it progresses from a point t_0 to a point t_1 in time and assigns it a value with unit [energy] $\times$ [time]. According to the Hamiltonian principle, the physical trajectory, *i.e.*, the actual field configuration that is realised in the evolution from t_0 to t_1, is determined by a stationary action, *i.e.*, when S assumes a minimal or maximal value. The Lagrangian density $\mathcal{L}$ incorporates the dynamics of the field ϕ and has a unit of energy.

As we shall see below, symmetries in physics can be studied systematically with the aid of Lie groups and their Lie algebras. Simply put, the *Lie algebra* $\mathfrak{g}$ of a *matrix Lie group* G is a vector space of all matrices X, such that $U(\alpha) = \mathrm{e}^{\mathrm{i}\alpha X} \in G$, together with an antisymmetric bilinear map (the "commutator" or "Lie bracket"),

$$[\,\cdot\,,\,\cdot\,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}\,, \quad [X, Y] = XY - YX\,, \tag{2.2}$$

which obeys the Jacobi identity:

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0 \,\forall\, X, Y, Z \in \mathfrak{g}\,. \tag{2.3}$$

The *generators* T^a of the Lie group G are a set $\{T^1, \ldots, T^n\}$, which forms a basis of the Lie algebra $\mathfrak{g}$. They obey the commutation relations

$$\left[T^a, T^b\right] = \mathrm{i} f^{abc} T^c\,. \tag{2.4}$$

Quantum Field Theory extends QM to *fields*, *i.e.*, reinterprets a field ϕ as an operator $\hat{\phi}$ (the hats are only used here as a distinction) on a suitable Hilbert space, imposing *canonical commutation relations* among them

$$\left[\hat{\phi}(t,\vec{x}), \hat{\phi}'(t,\vec{y})\right] = \mathrm{i}\delta^{(3)}(\vec{x}-\vec{y})\,, \tag{2.5}$$

$$\left[\hat{\phi}(t,\vec{x}), \hat{\phi}(t,\vec{y})\right] = \left[\hat{\phi}'(t,\vec{x}), \hat{\phi}'(t,\vec{y})\right] = 0\,. \tag{2.6}$$

The definition of the fields on a Hilbert space guarantees the existence of an inner product, which in turn guarantees that probability interpretation of the quantum fields is possible. Implementing commutation relations among the fields in a classical field theory *quantises* the theory, giving rise to so-called *creation* and *annihilation* operators, in terms of which the (then quantum) fields of the theory are expressed. As their names suggest, these operators have the simple interpretation to create and annihilate quantised excitations in the fields – the *particle states*.

Every quantum field is subject to the requirement that it transforms according to a representation of the Poincaré group, which represents the full set of continuous symmetries of special relativity, namely *translations*, *rotations*, and *boosts*. This requirement ensures that every QFT is manifestly relativistic, *i.e.*, is valid in high-energy and high-velocity regimes. Each representation has a characteristic intrinsic form of angular momentum, called *spin*. While elementary particles should not be thought of as spinning around their axes (the concept of an axis is ill-defined for point-like objects), the connection with angular momentum becomes apparent for instance in angular distributions in scattering experiments, as will be alluded to below.

The simplest representation of the Poincaré group is the scalar representation, which provides a spin-0 representation. A scalar field has a single component and, in the case

of a free complex field with mass m, its dynamics are entailed in the Lagrangian

$$\mathcal{L}_{\text{scalar}} = \frac{1}{2}\left(\partial_\mu \phi\right)^* \left(\partial^\mu \phi\right) - \frac{m^2}{2}\phi^* \phi \,. \tag{2.7}$$

For the case of SU(N)-type Lie groups, the construction of a local gauge theory is referred to as a Yang-Mills theory [58]. Let us consider a space-time-dependent set of parameters $\alpha(x) \equiv \{\alpha_i(x)\}$ and apply it as the following change of phase of a generic field Θ,

$$\Theta(x) \mapsto U(\alpha(x))\Theta(x)\,, \quad U(\alpha(x)) = \mathrm{e}^{\mathrm{i}\alpha_i(x)T_i}\,, \tag{2.8}$$

where the quantities T_i are the generators of the Lie group G. We now say that Θ transforms under the *fundamental* representation of G. Requiring the associated Lagrangian density to be invariant under the action of G introduces a vector (*i.e.* spin-1) field A^μ, called *gauge field*, which transforms under the *adjoint* representation of G,

$$A^\mu \mapsto U(\alpha(x))A^\mu U(\alpha(x))^{-1} + \frac{\mathrm{i}}{g}\left(\partial^\mu U(\alpha(x))\right) U(\alpha(x))^{-1}\,. \tag{2.9}$$

The invariance of the Lagrangian density is then ensured by replacing the derivative ∂_μ by the *gauge-covariant* derivative, defined as

$$D_\mu \Theta = \partial^\mu \cdot \mathbb{1} + \mathrm{i}g A_\mu \Theta\,. \tag{2.10}$$

We have introduced a new parameter g, which represents the *coupling constant* of the theory, coupling the field Θ to the gauge field A^μ.

As it transforms with respect to the adjoint representation of the group G, the gauge field A^μ is part of the Lie algebra of G and may therefore be written in terms of its generators, $A_\mu = A^a_\mu T^a$. Its dynamics are determined by the *field strength tensor* $F_{\mu\nu}\Theta = \left[D_\mu, D_\nu\right]\Theta$, in terms of which the free vector-field Lagrangian reads

$$\mathcal{L}_{\text{vector}} = -\frac{1}{4}F^{a,\mu\nu}F^a_{\mu\nu}\,, \tag{2.11}$$

where

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g f^{abc} A^b_\mu A^c_\nu\,. \tag{2.12}$$

In the case of an abelian group, the third term in the equation above vanishes and is absent in the Lagrangian.

For the discussion of the Standard Model, we shall also need spin-$\frac{1}{2}$ fields, corresponding to the spinor representation of the Poincaré group. Its Lagrangian density is given by

$$\mathcal{L}_{\text{spinor}} = \mathrm{i}\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi\,, \tag{2.13}$$

where we have introduced the Dirac matrices $\gamma^\mu = \{\gamma^0, \gamma^1, \gamma^2, \gamma^3\}$, a set of 4×4 matrices obeying the *anti-commutation* relation

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbb{1}_{4\times 4}\,, \tag{2.14}$$

where $g^{\mu\nu}$ denotes the Minkowski metric tensor.

2.2 The Standard Model

The Standard Model is constructed as a Yang-Mills gauge theory with gauge group given by the direct product of the three Lie groups

$$\mathrm{U}(1)_\mathrm{Y} \times \mathrm{SU}(2)_\mathrm{W} \times \mathrm{SU}(3)_\mathrm{C}\,. \tag{2.15}$$

Here, $\mathrm{SU}(3)_\mathrm{C}$ contributes the symmetry pertaining to the colour charges of quarks and gluons, *i.e.*, reflects Quantum Chromodynamics (QCD); the direct product $\mathrm{U}(1)_\mathrm{Y} \times \mathrm{SU}(2)_\mathrm{W}$ contributes electroweak (EW) interactions, including quantum electrodynamics (QED), weak gauge boson self-interactions, and the Higgs sector.

The fermionic matter in the SM is organised in six *doublets* over three *generations*, broken into the *quark* and *lepton* family,

$$\underbrace{\begin{pmatrix}\mathrm{u}\\ \mathrm{d}\end{pmatrix}, \begin{pmatrix}\mathrm{c}\\ \mathrm{s}\end{pmatrix}, \begin{pmatrix}\mathrm{t}\\ \mathrm{b}\end{pmatrix}}_{\text{Quarks}}, \quad \underbrace{\begin{pmatrix}\nu_\mathrm{e}\\ \mathrm{e}^-\end{pmatrix}, \begin{pmatrix}\nu_\mu\\ \mu^-\end{pmatrix}, \begin{pmatrix}\nu_\tau\\ \tau^-\end{pmatrix}}_{\text{Leptons}}\,. \tag{2.16}$$

We will write the SM Lagrangian as a sum over a QCD and an EW term,

$$\mathcal{L}_\mathrm{SM} = \mathcal{L}_\mathrm{QCD} + \mathcal{L}_\mathrm{EW}\,, \tag{2.17}$$

where the two terms $\mathcal{L}_\mathrm{QCD}$ and $\mathcal{L}_\mathrm{EW}$ will be given in section 2.2.1 and section 2.2.2 below, respectively.

2.2.1 Quantum Chromodynamics

The classical (*i.e.* non-quantised) gauge-invariant QCD Lagrangian is given by

$$\mathcal{L}^\mathrm{class}_\mathrm{QCD} = -\frac{1}{4}F^a_{\mu\nu}F^{a,\mu\nu} + \sum_{\substack{f\in\\ \{\mathrm{u,d,c,s,t,b}\}}} \bar{q}_{f,i}\left(\mathrm{i}\not{D}^i{}_j - m_f\delta^i{}_j\right)q^j_f\,, \tag{2.18}$$

where the gauge-covariant derivative is defined by

$$\not{D}^i{}_j = \gamma^\mu (D^\mu)^i{}_j = \gamma^\mu\left(\partial_\mu\cdot\mathbb{1} + \mathrm{i}g_\mathrm{S}A^a_\mu T^a\right)^i{}_j\,. \tag{2.19}$$

The classical Lagrangian eq. (2.18) can be quantised *e.g.* by the method of [59], which, in the so-called R_ξ gauge, introduces a *gauge-fixing* term,

$$\mathcal{L}^\mathrm{gf}_\mathrm{QCD} = -\frac{1}{2\xi}(\partial^\mu A^a_\mu)^2 \tag{2.20}$$

together with a *ghost field* Lagrangian

$$\mathcal{L}^\mathrm{ghost}_\mathrm{QCD} = \bar{\eta}^a\left(-\partial^\mu D^{ab}_\mu\right)\eta^b\,, \tag{2.21}$$

expressed in terms of the covariant derivative

$$D^{ab}_\mu = \delta^{ab}\partial_\mu + g_\mathrm{S}f^{abc}A^c_\mu\,. \tag{2.22}$$

The full quantised QCD Lagrangian is therefore given by

$$\mathcal{L}_\mathrm{QCD} = \mathcal{L}^\mathrm{class}_\mathrm{QCD} + \mathcal{L}^\mathrm{gf}_\mathrm{QCD} + \mathcal{L}^\mathrm{ghost}_\mathrm{QCD} \tag{2.23}$$

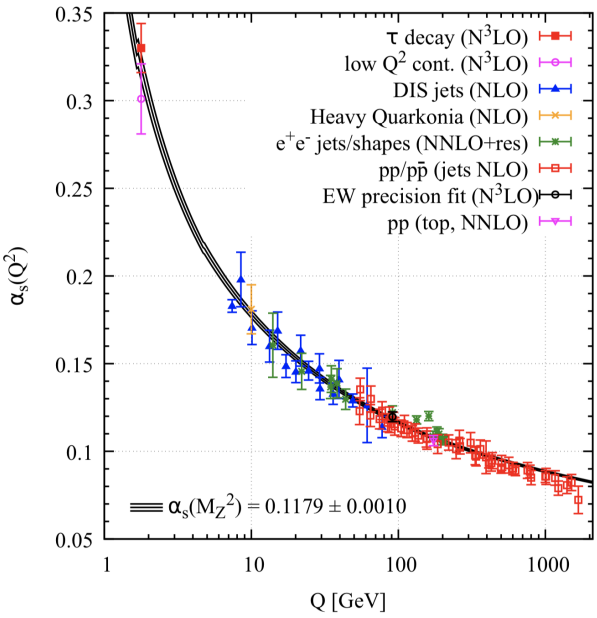
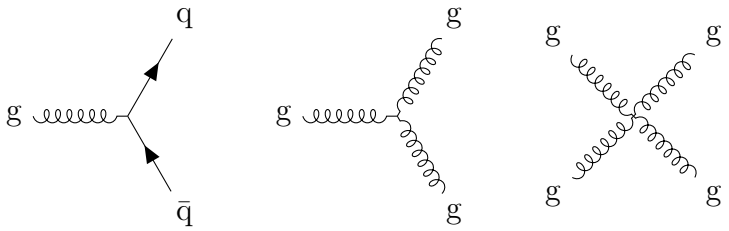


Figure 2.1: *Summary of measurements of the dependence of the strong coupling constant on the energy scale Q. The parentheses state the respective perturbative accuracy used in the extraction. Taken from* [60].

The terms in the classical part of the QCD Lagrangian give rise to the following three QCD vertices,



(2.24)

coupling gluons to a quark-antiquark pair, three gluons, and four gluons respectively.

Due to quantum corrections, the strong coupling constant $g_\mathrm{S} = \sqrt{4\pi\alpha_\mathrm{S}}$ varies with the energy scale, expressed in terms of the renormalisation group equation,

$$\mu_\mathrm{R}^2 \frac{\partial \alpha_\mathrm{S}(\mu_\mathrm{R}^2)}{\partial \mu_\mathrm{R}^2} = \beta(\alpha_\mathrm{S})\,. \tag{2.25}$$

Here, the auxiliary scale μ_R is called the *renormalisation scale* and the right-hand side of the equation denotes the *beta function* of QCD. Its most prominent feature is known as *asymptotic freedom* [61, 62], meaning that it drives the strong coupling α_S to become exceedingly large at small energy scales but smaller and smaller for higher and higher scales, *cf.* fig. 2.1. As a consequence, coloured particles are considered to behave like free particles at high energies, while being confined in bound states at low energies, a phenomenon known as *confinement*. The beta function can be expanded in a perturbative series

$$\beta(\alpha_\mathrm{S}) = -\sum_{i=0}^{\infty} \left(\frac{\beta_i}{4\pi}\right)^i \alpha_\mathrm{S}^{2+i}\,, \tag{2.26}$$

with the first two (renormalisation scheme independent) coefficients being given by

$$\beta_0 = \frac{11}{3} C_\mathrm{A} - \frac{4}{3} T_\mathrm{R} n_\mathrm{F}\,, \tag{2.27}$$

$$\beta_1 = \frac{34}{3}C_\mathrm{A}^2 - \frac{20}{3}C_\mathrm{A}T_\mathrm{R}n_\mathrm{F} - 4C_\mathrm{F}T_\mathrm{R}n_\mathrm{F}\,. \tag{2.28}$$

2.2.2 Electroweak Interactions

We decompose the classical electroweak Lagrangian into four terms,

$$\mathcal{L}_\mathrm{EW}^\mathrm{class} = \mathcal{L}_\mathrm{EW}^\mathrm{boson} + \mathcal{L}_\mathrm{EW}^\mathrm{fermion} + \mathcal{L}_\mathrm{EW}^\mathrm{Higgs} + \mathcal{L}_\mathrm{EW}^\mathrm{Yukawa}\,, \tag{2.29}$$

the bosonic part, fermionic part, Higgs sector, and the Yukawa interactions. To remove unphysical degrees of freedom, gauge-fixing and ghost terms similar to the ones in QCD, *cf.* eqs. (2.20) and (2.21), have to be added to the classical Lagrangian upon quantisation.

We will start with the bosonic piece,

$$\mathcal{L}_\mathrm{EW}^\mathrm{boson} = -\frac{1}{4}W_{\mu\nu}^a W^{a,\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}\,, \tag{2.30}$$

It determines the dynamics of the electroweak gauge bosons and consists of the field strength tensor

$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g\varepsilon^{abc}W_\mu^b W_\nu^c \tag{2.31}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.32}$$

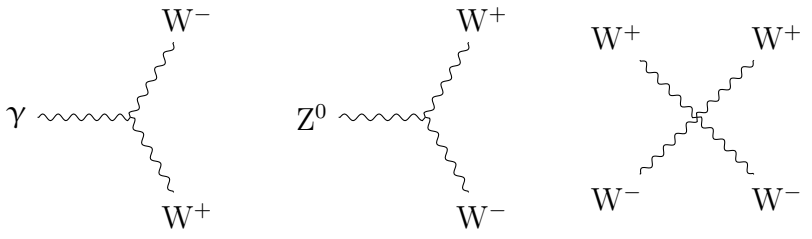
for the gauge boson B_μ of the (abelian) group $\mathrm{U}(1)_\mathrm{Y}$. The Levi-Civita symbol in the definition of $W_{\mu\nu}^a$ corresponds to the structure constants of $\mathfrak{su}(2)$, the Lie algebra of SU(2), given by the commutator

$$\left[\sigma^a, \sigma^b\right] = \mathrm{i}\varepsilon^{abc}\sigma^c \tag{2.33}$$

of the Pauli matrices σ^a, which act as the generators of the SU(2) Lie group. Following the Yang-Mills construction principle in section 2.1, the gauge fields W_μ^a and B_μ induce a gauge-covariant derivative,

$$D_\mu = \partial_\mu + \mathrm{i}g\frac{\sigma^a}{2}W_\mu^a + \mathrm{i}g'\frac{Y}{2}B_\mu\,, \tag{2.34}$$

where the second term corresponds to the $\mathrm{SU}(2)_\mathrm{W}$ group and the third term corresponds to the abelian $\mathrm{U}(1)_\mathrm{Y}$. Due to the non-abelian nature of the $\mathrm{SU}(2)_\mathrm{W}$ gauge group, the bosonic piece $\mathcal{L}_\mathrm{EW}^\mathrm{boson}$ introduces self-interaction vertices between the SU(2) gauge bosons, just as is the case for gluons in QCD,



(2.35)

The fermionic part of the EW Lagrangian is given by

$$\mathcal{L}_\mathrm{EW}^\mathrm{fermion} = \sum_{i=1}^{3}\mathrm{i}\bar{Q}_i \not{D} Q_i + \sum_{i=1}^{3}\mathrm{i}\bar{u}_i \not{D} u_i + \sum_{i=1}^{3}\mathrm{i}\bar{d}_i \not{D} d_i + \sum_{i=1}^{3}\mathrm{i}\bar{L}_i \not{D} L_i + \sum_{i=1}^{3}\mathrm{i}\bar{e}_i \not{D} e_i \tag{2.36}$$

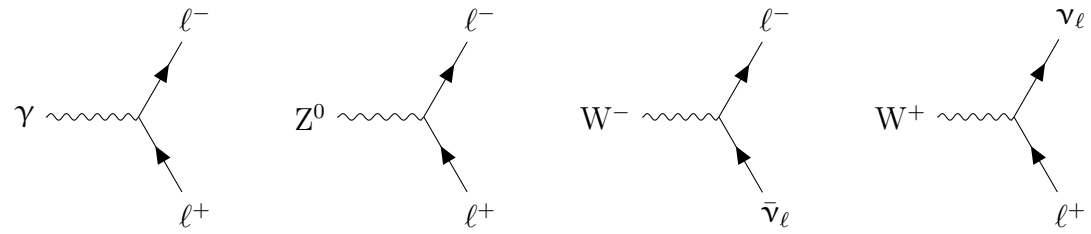
where

$$Q_i = \begin{pmatrix} u_i \\ d_i \end{pmatrix}_{\mathrm{L}} , \quad L_i = \begin{pmatrix} \nu_i \\ e_i \end{pmatrix}_{\mathrm{L}} \tag{2.37}$$

denote the left-handed quark and fermion field doublets, respectively, and

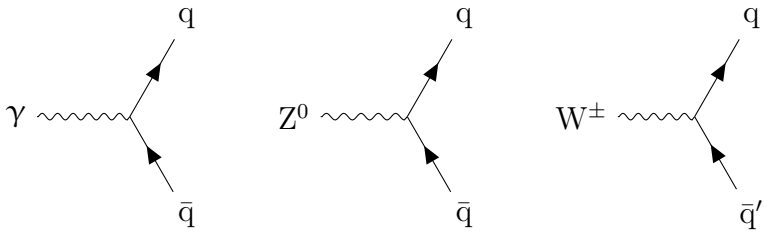
$$u_i \equiv (u_i)_{\mathrm{R}} , \quad d_i \equiv (d_i)_{\mathrm{R}} , \quad e_i \equiv (e_i)_{\mathrm{R}} \tag{2.38}$$

denote the right-handed up-type, down-type, and charged-fermion field singlets, respectively. In this notation, the corresponding fields of the three particle generations are simply enumerated, such that u_1 denotes the up-quark field, d_2 the strange-quark field, e_3 the τ^--lepton field, etc. The fermionic piece of the EW Lagrangian introduces vertices between the gauge bosons and leptons,



(2.39)

and gauge bosons and quarks,

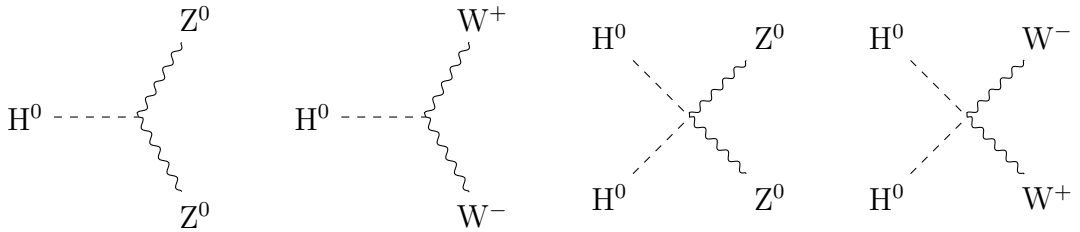


(2.40)

We will now turn to the Higgs piece of the EW Lagrangian. Since it is a scalar field, the Higgs contributes the simplest piece,

$$\mathcal{L}_{\mathrm{EW}}^{\mathrm{Higgs}} = \left(D_\mu \phi\right)^\dagger \left(D^\mu \phi\right) - \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2 . \tag{2.41}$$

It introduces triple and quartic couplings between the Higgs and the massive gauge bosons,

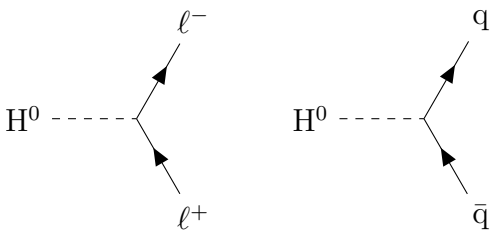


(2.42)

Closely connected to the Higgs piece is the Yukawa piece of the EW Lagrangian,

$$\mathcal{L}_{\mathrm{EW}}^{\mathrm{Yukawa}} = -\sum_{i,j=1}^{3} y_{\mathrm{u},ij} \bar{Q}_i (\mathrm{i}\sigma^2 \phi^\dagger) u_j - \sum_{i,j=1}^{3} y_{\mathrm{d},ij} \bar{Q}_i \phi^\dagger d_j - \sum_{i,j=1}^{3} y_{\ell,ij} \bar{L}_i \phi^\dagger e_j + \mathrm{h.c.} \ , \tag{2.43}$$

which introduces couplings between the scalar field ϕ and (massive) fermions,



(2.44)

Note, that due to the absence of right-handed neutrinos, *cf. e.g.* [63], the mass term for neutrinos is absent in eq. (2.43).

The masses of the electroweak gauge bosons and fermions are generated in a gauge-invariant way via the *Higgs mechanism*[1] [64–66]. If the parameter μ^2 in eq. (2.41) is chosen to be negative, the Higgs potential has an infinite set of minima for

$$\phi^\dagger\phi = -\frac{\mu^2}{2\lambda}\,. \tag{2.45}$$

The EW symmetry can therefore be broken *spontaneously*, meaning, that it exists on the level of the Lagrangian, but is broken by choosing one of the degenerate vacuum states. As the vacuum must not be electrically charged, the minimum of the Higgs field is chosen to be

$$\langle\phi\rangle_0 = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix}\,, \tag{2.46}$$

with the *vacuum expectation value* (vev) v. Expanding the Higgs field around its vacuum state

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v+h(x)\end{pmatrix}\,, \tag{2.47}$$

generates mass matrices for the gauge-boson fields W^a_μ and B_μ through the covariant derivative D_μ in eq. (2.41). Diagonalisation yields the four new fields

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu \pm W^2_\mu\right)\,, \tag{2.48}$$

$$Z_\mu = \frac{1}{\sqrt{g^2+g'^2}}\left(gW^3_\mu - g'B_\mu\right)\,, \tag{2.49}$$

$$A_\mu = \frac{1}{\sqrt{g^2+g'^2}}\left(g'W^3_\mu + gB_\mu\right)\,, \tag{2.50}$$

corresponding to the charged $\mathrm{W}^\pm$ bosons, the neutral Z^0 boson, and the photon γ, respectively. The former acquire masses

$$m_{\mathrm{W}^\pm} = \frac{v}{2}g\,, m_{\mathrm{Z}^0} = \frac{v}{2}\sqrt{g^2+g'^2}\,, \tag{2.51}$$

while the latter remains massless, $m_\gamma = 0$. An equivalent procedure generates the mass matrices of the fermions, which in the case of the quarks is known as the CKM [2] matrix [67, 68] and in case of the leptons as the PMNS [3] matrix [69, 70].

[1] To be precise it should be called the *Englert-Brout-Higgs-Guralnik-Hagen-Kibble mechanism.*
[2] Cabibbo-Kobayashi-Maskawa.
[3] Pontecorvo–Maki–Nakagawa–Sakata.

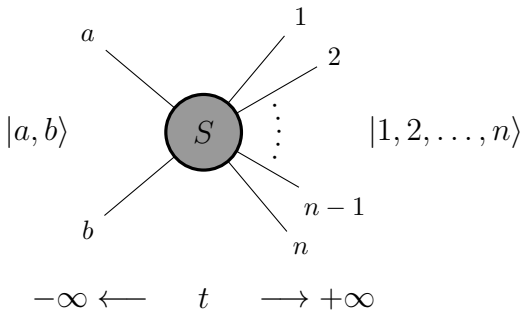


Figure 2.2: *Illustration of the scattering matrix S.*

2.3 From Lagrangians to Cross Sections

Having QFT and the Feynman rules of the Standard Model at hand, we are now in a position to connect the fundamental theory with observable phenomena. As the most basic quantity of interest, we may ask for the probability of a certain reaction $a + b \to n$ to happen, *i.e.*, for the probability that a final state comprised of the particles $1, 2, \ldots, n$ is created by the interaction of two initial-state particles a and b. Taking the quantum nature of fundamental particles into account, we are interested in calculating the transition rate for a quantum state $|a, b\rangle$ to evolve into a quantum state $|1, 2, \ldots, n\rangle$, which we can express via a *scattering matrix* $\mathcal{M}_{ab\to n}$ [71, 72]

$$\mathcal{M}_{ab\to n} = \langle 1, 2, \ldots, n| \, S \, |a, b\rangle \ . \tag{2.52}$$

Here, the interaction taking place in the process $ab \to n$ is described by the operator S, while $|a, b\rangle$ and $|1, 2, \ldots, n\rangle$ are (non-interacting) asymptotically free states in the distant past $t \to -\infty$ and distant future $t \to +\infty$ [4]. As a consequence of conservation of probability in quantum mechanics, the operator S has to be *unitary*. By application of time-ordered perturbation theory, the scattering matrix elements $\mathcal{M}_{ab\to n}$ can be directly related to the Feynman rules of the given QFT, giving a shortcut to the calculation of the scattering matrix elements. A derivation of the Standard-Model Feynman rules from time-ordered perturbation theory can be found *e.g.* in [57].

The *cross section* $\hat{\sigma}_{ab\to n}$ of the process $ab \to n$ is calculated by integrating the matrix element over the available phase space and normalising it by the flux of initial-state particles,

$$\hat{\sigma}_{ab\to n} = \frac{1}{4\sqrt{(p_a \cdot p_b)^2 - m_a^2 m_b^2}} \int \mathrm{d}\Phi_n \, |\mathcal{M}_{ab\to n}|^2 \, , \tag{2.53}$$

where the n-particle Lorentz-invariant phase space measure is given by

$$\mathrm{d}\Phi_n \equiv \mathrm{d}\Phi_n(p_a + p_b; p_1, \ldots, p_n) = (2\pi)^4 \delta^{(4)} \left(p_a + p_b - \sum_{i=1}^{n} p_i \right) \prod_{i=1}^{n} \frac{\mathrm{d}^4 p_i}{(2\pi)^3} \delta(p_i^2 - m_i^2) \, . \tag{2.54}$$

For ideal point-like particles a and b eq. (2.53) would already provide the full answer to the question after the interaction rate of $ab \to n$. Often, however, we are interested in collisions where either one (deeply-inelastic scattering – DIS) or both (hadron-hadron collisions) of the initial-state particles are hadrons, which intrinsically are extended objects. Moreover, even point-like particles can behave like composite objects due to radiative corrections. To give an example, an electron may radiate a photon, which in turn takes

[4] Mathematically, a decomposition of an "interaction state" into a direct product of one-particle states is not possible, so that the in- and out-states have to be considered asymptotically free. Physically, this implies that there is no particle interpretation of a scattering state.

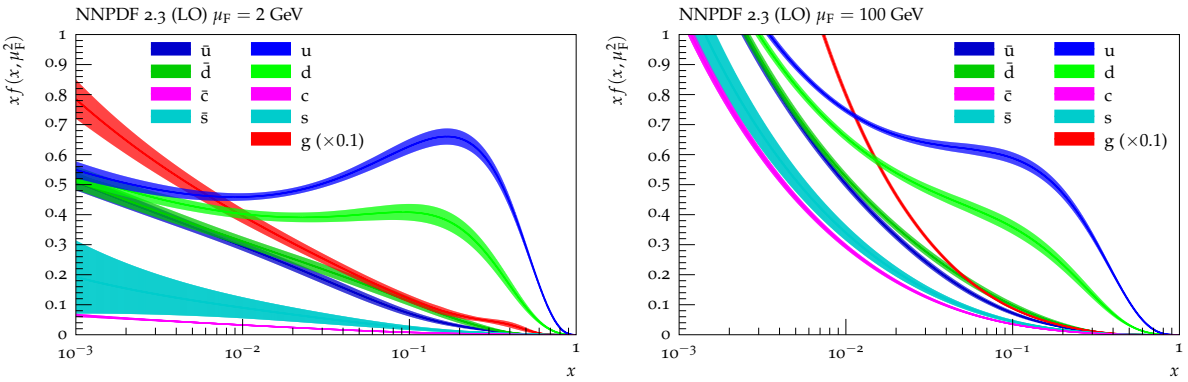


Figure 2.3: *Visualisation of proton PDFs at hand of the NNPDF 2.3 leading-order PDF set* [80] *evaluated at* $\mu_F = 2$ GeV *(left) and* $\mu_F = 100$ GeV *(right). Note that the gluon PDFs are divided by a factor 10. Generated with the* APFELweb *tool* [81, 82] *using the* LHAPDF6 *PDF library* [83] *and* Rivet [84, 85].

part in the hard interaction; this could then be interpreted as a photon from "inside" the electron.

Because cross sections are a combination of long- and short-distance phenomena. Following [73–79], we can obtain the *hadronic* cross section σ from the *partonic* cross section $\hat{\sigma}$ by means of factorisation, *i.e.*, by assuming that the scattering takes place between two quasi-free partons which do not depend on the other partons inside the hadron. This yields the following master formula for the *inclusive* cross section

$$\sigma_{h_1 h_2 \to n} = \sum_{a,b} \int_0^1 \int_0^1 \mathrm{d}x_a \, \mathrm{d}x_b \, f_{a/h_1}(x_a, \mu_F^2) f_{b/h_2}(x_b, \mu_F^2) \hat{\sigma}_{ab\to n}(\mu_F^2, \mu_R^2) \,, \tag{2.55}$$

where we have introduced the *parton distribution functions* (PDFs) f_{a/h_1} and f_{b/h_2} and the *factorisation scale* μ_F. (The renormalisation scale μ_R was already introduced in section 2.2.) Broadly speaking, the functions f_{a/h_1} and f_{b/h_2} represent the probability to find a parton a in the hadron h_1 with a fraction x_a of its momentum, $p_a = x_a p_{h_1}$, at the energy scale μ_F (a parton b in hadron h_2 with fraction x_b of its momentum, $p_b = x_b p_{h_2}$ at scale μ_F). These functions account for the composite nature of hadrons. The sum over a and b in eq. (2.55) takes into account all possible *partonic channels*, *i.e.*, possible initial states that can be formed from the hadronic constituents and yield the respective final state.

Parton distribution functions have historically been introduced and measured in deeply inelastic scattering (DIS), *i.e.*, highly energetic lepton-hadron collisions. While they cannot be computed from first principles, perturbative QCD gives a recipe to evolve PDFs in the factorisation scale μ_F. Taking into account that partons may branch according to the Feynman rules of QCD, *cf.* section 2.2.1, the scaling behaviour of PDFs is encapsulated in an equation similar in spirit to the renormalisation group equation eq. (2.25),

$$\mu_F^2 \frac{\mathrm{d} f_{j/h}(x, \mu_F^2)}{\mathrm{d}\mu_F^2} = \frac{\alpha_S(\mu_R^2)}{2\pi} \sum_{i=q,g} \int_x^1 \frac{\mathrm{d}\xi}{\xi} f_{i/h}(\xi, \mu_F^2) P_{ij}\left(\frac{x}{\xi}\right) \,, \tag{2.56}$$

known as the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equation [73, 86, 87]. It has to be read as follows: the change of the probability to find a parton of flavour j in hadron h with the energy scale μ_F is governed by the sum of all ways this flavour can be

produced in a branching $i \mapsto jk$ weighted by the probability[5] of this branching to occur, represented by the splitting kernels P_{ij}, *cf.* eqs. (3.59) to (3.61). These splitting kernels and DGLAP evolution will be discussed in more detail in section 3.2.1. An illustration of the scale evolution of proton PDFs is given in fig. 2.3. It is interesting to note that the "valence quark" contributions to the proton peak at larger momentum fractions x, while the bulk of the proton constituents is given by "sea quarks" at low x. This is in alignment with the naive proton picture, because it means that a parton with large momentum fraction is most likely to be an up or down quark. A parton with small momentum fraction is more likely to be produced by a branching inside the hadron.

Although collinear factorisation builds upon a rigorous proof only for DIS and Drell-Yan processes [73–79], the "divide-and-conquer" approach to factorise the full (hadronic) process into partonic and hadronic contributions proved itself very successful, even for processes where the validity has not yet been shown. The determination of factorisation-breaking contributions, *cf. e.g.* [88], as well as so-called higher-twist contributions, *i.e.*, corrections beyond the leading power which allow reactions with multiple partons simultaneously, *cf. e.g.* [89], remains an active field of research. Especially the latter becomes relevant in the discussion of higher-order corrections to the leading-twist contribution, as they may be of similar importance.

Mostly, we are interested in *differential* cross sections, *i.e.*, the rate to observe a certain reaction while also measuring further quantities, such as the particle momenta or angles. Denoting the observable of interest by O, the differential cross section is given by

$$\begin{aligned}\frac{\mathrm{d}\sigma_{h_1h_2\to n}}{\mathrm{d}O} &= \sum_{a,b} \int_0^1 \int_0^1 \mathrm{d}x_a\, \mathrm{d}x_b\, f_{a/h_1}(x_a,\mu_\mathrm{F}^2) f_{b/h_2}(x_b,\mu_\mathrm{F}^2) \frac{\mathrm{d}\hat{\sigma}_{ab\to n}(\mu_\mathrm{F}^2,\mu_\mathrm{R}^2)}{\mathrm{d}O} \\ &= \frac{1}{2s} \sum_{a,b} \int_0^1 \int_0^1 \frac{\mathrm{d}x_a}{x_a} \frac{\mathrm{d}x_b}{x_b}\, f_{a/h_1}(x_a,\mu_\mathrm{F}^2) f_{b/h_2}(x_b,\mu_\mathrm{F}^2) \\ &\qquad \times \int \mathrm{d}\Phi_n\, |\mathcal{M}_{ab\to X}|^2(\Phi_n;\mu_\mathrm{F}^2,\mu_\mathrm{R}^2)\, \delta\left(O - O(\Phi_n)\right) \,, \end{aligned} \tag{2.57}$$

where we have used an approximated flux factor in the limit of massless initial-state particles,

$$F = \frac{1}{4\sqrt{(p_a\cdot p_b)^2 - m_a^2 m_b^2}} \xrightarrow{m_a\to 0, m_b\to 0} \frac{1}{4 p_a\cdot p_b} = \frac{1}{2s}\frac{1}{x_a x_b} \,, \tag{2.58}$$

and inserted the light-cone momentum fractions to express the result in terms of the centre-of-mass energy squared $s = E_\mathrm{CM}^2$

$$2p_a\cdot p_b = 2(x_a p_{h_1})\cdot(x_b p_{h_2}) = 2p_{h_1}\cdot p_{h_2} x_a x_b = x_a x_b s \,.$$

[5] Strictly speaking, a probability interpretation of the DGLAP kernels is only possible at leading order.

3

Monte Carlo Event Generators

Monte Carlo event generators [90] bridge the gap between the mathematically challenging Quantum Field Theory underlying modern particle physics and measurements at collider experiments. Multi-purpose event-generation frameworks such as PYTHIA [91–102], HERWIG [103–109], and SHERPA [110–112] and more specialised frameworks like WHIZARD [113] and MADGRAPH_AMC@NLO [114, 115] provide *exclusive* descriptions of the complex events produced in high-energy particle collisions. This means, an event generator will produce a list of particles, an *event record*, just as it would have been measured in a (perfect) detector, including their momenta, charges, and spins. As such, they provide the means to describe and test our understanding of particle physics and as such our understanding of the fundamental laws of physics.

This chapter firstly provides an overview of the methodology employed in particle physics phenomenology, both theoretically and in terms of algorithmic implementations. Secondly, an attempt is made to give a fair account of the existing literature, where relevant. As such, this chapter shall put the research presented in the later chapters of this thesis into broader perspective.

After an overview of the description of the hard scattering, particular emphasis will be placed on the modelling of the perturbative cascade from the high energy scale of the hard process down to the low energy scale at which experimentally measurable particles emerge. Within this, numerical as well as analytical methods will be discussed, building the foundation for the core theme of this thesis. Subsequently, the systematic improvement of these approximations by fixed-order corrections (so-called matching and merging schemes) will be discussed in similar detail. This chapter will be closed by a summary of hadronisation models, the last step in the process of event generation.

Most of the theoretical basics discussed follow closely [116] and [117], while parts refer to the introductions in [118, 119].

3.1 The Hard Process

For processes that involve a hard scale, the event generation commences with the *hard process*, *i.e.*, the basic interaction of two beams producing a (small) number of outgoing particles. To this end, the inclusive cross section $\sigma_{ab\to n}$ has to be calculated according to eq. (2.55). This is done by means of perturbation theory, which means that contributions to the matrix element $\mathcal{M}$ are organised in a series expansion of the coupling constants,

$g_S = \sqrt{4\pi\alpha_S}$ and $e = \sqrt{4\pi\alpha}$,

$$\begin{aligned}\mathcal{M}_{ab\to n} &= \sum_{i=0}^{\infty}\sum_{k,\ell=0}^{\infty}\mathcal{M}_{n+i}^{(k,\ell)} \\ &= \mathcal{M}_n^{(0,0)} + g_S\mathcal{M}_{n+1}^{(0,0)} + g_S^2\left(\mathcal{M}_n^{(1,0)} + \mathcal{M}_{n+2}^{(0,0)}\right) \\ &\qquad + e\tilde{\mathcal{M}}_{n+1}^{(0,0)} + e^2\left(\mathcal{M}_n^{(0,1)} + \tilde{\mathcal{M}}_{n+2}^{(0,0)}\right) + \dots\end{aligned} \tag{3.1}$$

where lower indices represent additional legs due to QCD and QED/EW emissions (*i.e.*, *real* contributions), while upper indices represent additional QCD and QED/EW loops in the diagram (*i.e.*, *virtual* contributions). Note that each coefficient $\mathcal{M}_{n+i}^{(k,\ell)}$ contains a fixed number of additional, process-dependent powers of e and g_S, not made explicit in the expansion. On the level of the squared matrix element, $|\mathcal{M}_{ab\to n}|^2$, the squared $n+1$-particle tree-level amplitude $\left|\mathcal{M}_{n+1}^{(0,0)}\right|^2$ enters at the same order g_S^2 as the n-particle tree-one-loop interference $2\,\mathrm{Re}\left[\mathcal{M}_n^{(0,0)}\mathcal{M}_n^{(1,0)*}\right]$ (similarly for the QED/EW contributions). The square of the $n+2$-particle contribution, on the other hand, enters only at order g_S^4.

At *leading order* (LO), only the first term in the series above contributes, so that the cross section reads

$$\sigma_n^{\mathrm{LO}} = \frac{1}{2s}\sum_{a,b}\int_0^1\int_0^1 \frac{\mathrm{d}x_a}{x_a}\frac{\mathrm{d}x_b}{x_b}\, f_{a/h_1}(x_a,\mu_F^2) f_{b/h_2}(x_b,\mu_F^2)\int \mathrm{d}\Phi_n \left|\mathcal{M}_n^{(0,0)}\right|^2(\Phi_n;\mu_F^2,\mu_R^2)\,. \tag{3.2}$$

A numerical implementation of this consists of two parts: the implementation of tree-level matrix elements (*i.e.*, ones without internal loops) and a phase-space generator, randomly sampling phase-space points to evaluate the matrix element on. The latter will be the subject of section 3.1.1, while matrix-element calculations are discussed in section 3.1.2. In the following, we will focus on higher-order QCD corrections and use the notation $\mathcal{M}_n^{(\ell)} \equiv \mathcal{M}_n^{(\ell,0)}$, which implicitly amounts to the lowest-order in QED/EW, to avoid clutter.

At next-to-leading order (NLO), *i.e.*, to first order in the coupling beyond the leading (*Born*) contribution, two new types of contributions enter the cross section: Born-one-loop interferences and squared $n+1$ particle amplitudes. Restricting ourselves to the case of QCD, the NLO cross section thus reads

$$\sigma_n^{\mathrm{NLO}} = \int \mathrm{d}\Phi_B\left[\mathrm{B}_n(\Phi_B;\mu_R,\mu_F) + \mathrm{V}_n(\Phi_B;\mu_R,\mu_F)\right] + \int \mathrm{d}\Phi_R\,\mathrm{R}_n(\Phi_R;\mu_R,\mu_F)\,, \tag{3.3}$$

where we have introduced the notation

$$\mathrm{B}_n(\Phi_B;\mu_R,\mu_F) := \left|\mathcal{M}_n^{(0)}\right|^2(\Phi_B;\mu_R,\mu_F)\,, \tag{3.4}$$

$$\mathrm{V}_n(\Phi_B;\mu_R,\mu_F) := 2\,\mathrm{Re}\left[\mathcal{M}_n^{(0)}\mathcal{M}_n^{(1)*}\right](\Phi_B;\mu_R,\mu_F)\,, \tag{3.5}$$

$$\mathrm{R}_n(\Phi_R;\mu_R,\mu_F) := \left|\mathcal{M}_{n+1}^{(0)}\right|^2(\Phi_R;\mu_R,\mu_F) \equiv \mathrm{B}_{n+1}(\Phi_B;\mu_R,\mu_F)\,. \tag{3.6}$$

and

$$\mathrm{d}\Phi_B := \frac{1}{2s}\sum_{A,B}\frac{\mathrm{d}x_A}{x_A}\frac{\mathrm{d}x_B}{x_B}\, f_{A/h_1}(x_A,\mu_F^2) f_{B/h_2}(x_B,\mu_F^2)\,\mathrm{d}\Phi_n, \tag{3.7}$$

$$\mathrm{d}\Phi_{\mathrm{R}} := \frac{1}{2s}\sum_{a,b}\frac{\mathrm{d}x_a}{x_a}\frac{\mathrm{d}x_b}{x_b}\, f_{a/h_1}(x_a,\mu_{\mathrm{F}}^2) f_{b/h_2}(x_b,\mu_{\mathrm{F}}^2)\,\mathrm{d}\Phi_{n+1}\,. \tag{3.8}$$

In addition to tree-level matrix elements for the Born and Born + 1-parton state, an NLO calculation requires a one-loop matrix element for the Born state, *cf.* section 3.1.3. Because the virtual and real corrections in eq. (3.3) are separately divergent, a numerical implementation necessitates the removal of these divergences. Techniques to do this will be discussed in section 3.1.4.

Today, the calculation of cross sections and observables at NLO in QCD defines the state of the art. It is a core part of the MadGraph_aMc@Nlo [114,115], Whizard [113], Matchbox [120, 121], and Sherpa [110–112] frameworks as well as dedicated particle-level Monte Carlo codes such as Mcfm [122–124], Helac-Nlo [125], or Vbfnlo [126–128], which is specialised on vector boson fusion processes and their backgrounds. The MoCaNlo + Recola framework [129–133] provides calculations of electroweak processes, including both NLO QCD and EW corrections.

For most $2 \to 2$ and $2 \to 3$ processes it is also possible to calculate differential cross sections up to the *next-to-next-to-leading order* (NNLO) in QCD, publicly available for some processes[1], *e.g.* in the public codes Matrix [134] and Mcfm starting from version 8 [135, 136]. The Nnlojet parton-level Monte Carlo generator [137–140] offers a wide variety of processes at NNLO, facilitated by the use of the antenna subtraction formalism, *cf.* section 3.1.4 below.

Most general-purpose Monte Carlo frameworks offer a selection of internal hard processes at LO and NLO. Dedicated hard-process Monte Carlos, on the other hand, often do not offer descriptions of additional logarithmically enhanced radiation[2] and non-perturbative hadronisation models. The Les Houches accord [143, 144] defines a format to exchange "event records", *i.e.*, fully-differential lists of particles, between different Monte Carlo programs, either in plain-text files or, more recently, via binary HDF5 event files [145].

3.1.1 Monte Carlo Phase Space

Apart from the generation (or implementation) of matrix elements, a core part of hard-process generators is the phase-space sampling, for which Monte Carlo methods are employed. Useful introductions can for instance be found in [146, 147].

Monte Carlo integration builds upon the idea that the (hyper-)volume of an object can be calculated by knowing an enveloping region and randomly sampling points inside this area, counting the number of times a point landed inside the object. Formulating this mathematically, we obtain

$$\langle I(f)\rangle_x = \frac{V(\Omega)}{N}\sum_{i=1}^{N} f(\vec{x}_i) \xrightarrow{N\to\infty} I = \int_\Omega \mathrm{d}^D\vec{x}\, f(\vec{x})\,, \tag{3.9}$$

for an integrable function f. The convergence of the Monte Carlo estimator $\langle I(f)\rangle_x$ to the true integral value I, indicated by the arrow in the equation above, is a consequence of the law of large numbers. Assuming the function f is square-integrable, we can introduce

[1] Other codes (*e.g.* Sherpa) also offer limited NNLO functionality for a small number of processes.

[2] Some generators implement analytical resummation techniques, such as CuTe-Mcfm [141] or Matrix + Radish [142].

the Monte Carlo estimate of the variance $\sigma^2(f)$,

$$\left\langle \sigma^2(f) \right\rangle_x = \frac{1}{N}\sum_{i=1}^{N} f^2(\vec{x}_i) - \left(\frac{1}{N}\sum_{i=1}^{N} f(\vec{x}_i)\right)^2 = \left\langle I(f^2) \right\rangle_x - \left\langle I(f) \right\rangle_x , \tag{3.10}$$

which gives a probabilistic estimation of the integration error. The above shows that independently of the dimensionality of the phase space the integration error scales as $\sim 1/\sqrt{N}$ in the Monte Carlo method, which builds the foundation of its application in event generators. While Monte Carlo integration is feasible for any integrable function, the error estimate is only viable if the function is also square-integrable.

Simple and well-known techniques build up the phase space recursively by exploiting the factorisation property of the phase-space measure [148, 149],

$$\mathrm{d}\Phi_n(p_a{+}p_b; p_1, \ldots, p_n) = \mathrm{d}\Phi_{n-m+1}(p_a{+}p_b; p_1, \ldots, p_{n-m+1}, q)\frac{\mathrm{d}q^2}{2\pi}\,\mathrm{d}\Phi_m(q; p_{n-m+1}, \ldots, p_n) , \tag{3.11}$$

with a pseudo-particle q of mass q^2. This allows to reduce an n-particle phase-space measure to an iteration of two-particle phase spaces,

$$\mathrm{d}\Phi_2(Q; p_i, p_j) = \frac{\lambda(s_{ij}, m_i^2, m_j^2)}{16\pi\, s_{ij}}\,\mathrm{d}\cos\theta_i \frac{\mathrm{d}\phi_i}{2\pi} , \tag{3.12}$$

with the Källén function $\lambda(a, b, c) = a^2{+}b^2{+}c^2{-}2ab{-}2ac{-}2bc$. This type of phase-space generator is known as an M-generator, since massive pseudo-particles are inserted. A fast and (for massless particles) flat phase-space generation algorithm is given by RAMBO [150, 151], which samples the n-particle phase space isotropically. Since it is agnostic to any kind of structure of squared matrix elements, RAMBO is mainly useful for testing purposes and does not provide a competitive phase-space integrator.

Since phase-space integration is one of the major efficiency bottlenecks in the generation of the hard-scattering cross section, a lot of work has gone into improving the sampling according the complex structure of matrix elements. Typical strategies can be broadly divided into two classes: *importance sampling* and *stratified sampling*. The former samples phase-space points according to a distribution close to the squared matrix element, while the latter divides the integration domain into a series of bins and samples points according to different distributions in each bin. Examples of the former are the SARGE [152] and HAAG [153] algorithms, which are well-suited for pure-QCD amplitudes, as they distribute points according to the leading singularities of QCD antenna functions (*cf.* section 3.1.4 and section 3.2.3). Well-known examples of the latter are adaptive techniques, such as VEGAS [154, 155] and the metropolis algorithm [156, 157], which facilitate the integration by dividing the domain into regions according to their relative importance. The above strategies have been automated in a series of implementations, *cf. e.g.* [158–162], possibly employing parallelisation techniques [124, 163].

Interesting alternatives are forward-branching phase space generators [164–168], which start from the Born-level phase and generate the phase space of additional radiation by successive momentum-conserving on-shell $2 \mapsto 3$ mappings, akin to strategies routinely employed in parton showers, *cf.* section 3.2. Such techniques have the strength that they allow for unweighted NLO event generation, as every Born configuration can be assigned a local NLO weight ("Born-local K-factor"). Recently, there has also been tremendous interest in techniques utilising neural networks and machine learning, *cf. e.g.* [169–176].

3.1.2 Tree-Level Matrix Elements

Feynman diagrams without internal loops are called *tree-level* diagrams and appear at any perturbative order. Examples of $2 \to 2$ and $2 \to 3$ tree-level diagrams are given in fig. 3.1. To facilitate efficient calculations of tree-level QCD matrix elements, it is customary to make the colour factors, helicity dependence, and kinematical part explicit in the amplitudes. Mathematically, this corresponds to decomposing the amplitude over a (possibly overcomplete) spanning set,

$$\mathcal{M}_n^{(0)}(\{h_i\},\{c_i\}) = \left\langle c_1, c_2, \ldots, c_n; h_1, h_2, \ldots, h_n \middle| \mathcal{M}_n^{(0)} \right\rangle , \tag{3.13}$$

so that the projection onto the colour basis yields *colour-ordered amplitudes* [177],

$$\left\langle c_1, c_2, \ldots, c_n \middle| \mathcal{M}_n^{(0)} \right\rangle = \mathcal{A}_n^{(0)}(p_1, p_2, \ldots, p_n) . \tag{3.14}$$

and the projection onto the helicity basis yields *helicity amplitudes*, *cf. e.g.* [178] for a review,

$$\left\langle c_1, c_2, \ldots, c_n; h_1, h_2, \ldots, h_n \middle| \mathcal{M}_n^{(0)} \right\rangle = \mathcal{A}_n^{(0)}(p_1^{h_1}, p_2^{h_2}, \ldots, p_n^{h_n}) , \tag{3.15}$$

Unless helicities are made explicit via superscripts, we will assume $\mathcal{A}_n$ to denote helicity-averaged amplitudes.

In the case of a pure-gluon amplitude, the *trace basis* provides a simple spanning set of the colour space,

$$\mathcal{M}_n^{(0)} = \sum_{\sigma \in S_{n-1}} \mathrm{Tr}(T^{a_1} T^{a_2} \ldots T^{a_n}) \mathcal{A}_n^{(0)}(p_1, p_{\sigma(2)}, \ldots, p_{\sigma(n)}) , \tag{3.16}$$

For the implementation in event generators, the *colour-flow basis* is especially useful, because it assigns each colour-ordered amplitude a distinct colour topology in terms of fundamental and antifundamental indices ("quark lines") [179–181]. Again for the case of a pure-gluon amplitude,

$$\mathcal{M}_n^{(0)} = \sum_{\sigma \in S_{n-1}} \delta^{j_{\sigma(2)}}{}_{i_1} \delta^{j_{\sigma(3)}}{}_{i_{\sigma(2)}} \cdots \delta^{j_1}{}_{i_{\sigma(n)}} \mathcal{A}_n^{(0)}(p_1, p_{\sigma(2)}, \ldots, p_{\sigma(n)}) . \tag{3.17}$$

The trace basis and colour-flow basis have the same dimensionality[3] and can therefore be related to each other. The basis transformation is given by the Fierz identity,

$$(T^a)^i{}_j (T^a)^k{}_\ell = T_\mathrm{R} \left(\delta^i{}_\ell \delta^k{}_j - \frac{1}{N_\mathrm{C}} \delta^i{}_j \delta^k{}_\ell \right) \tag{3.18}$$

The most direct way to implement tree-level matrix elements in Monte Carlo codes is by explicitly writing analytically calculated amplitudes into the source code. This strategy is pursued for the internal matrix-element libraries in PYTHIA and HERWIG, as well as the ALPGEN parton-level event generator [182]. Since this procedure quickly becomes cumbersome and the construction and evaluation of Feynman diagrams follows simple enough rules, it is possible to automate the calculation of tree-level matrix elements. This can either be done by explicit Feynman-diagram generation as done in AMEGIC++ [183] or MADGRAPH [115, 184, 185] and O'MEGA [113, 186], which use (variants of) the HELAS/HELAC algorithm [179, 187] to calculate wave functions for internal and external particle lines, which are later squared. More efficient algorithms [188], such as the one implemented in COMIX [161], are based on recursion relations [189, 190], which recursively

[3] Which is a non-trivial statement, because both "bases" are overcomplete spanning sets.

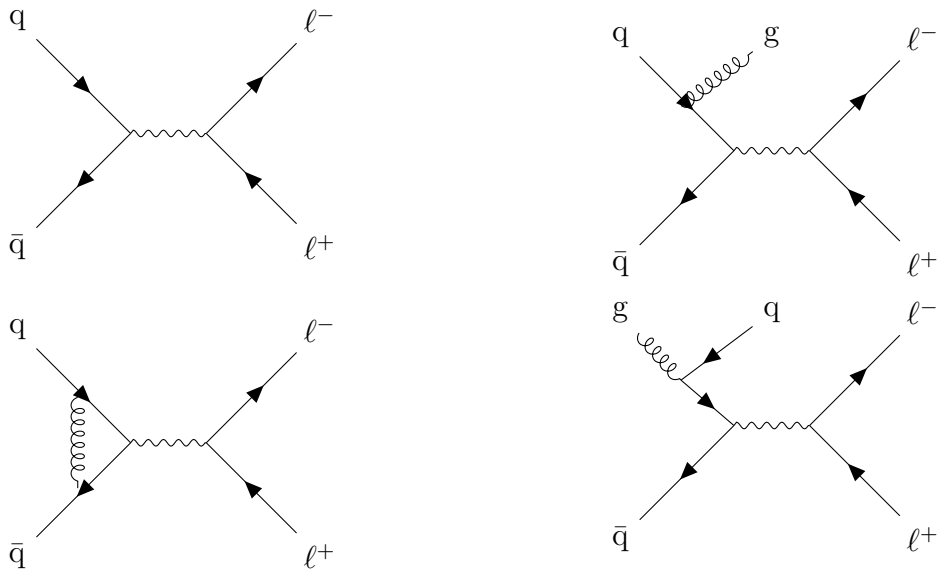


Figure 3.1: *Born (top left), virtual correction (bottom left), and real-radiation corrections (right) of the Drell-Yan process* $pp \to \ell^+\ell^-$.

relate so-called "off-shell currents" with decreasing number of legs, In this context, an off-shell current is defined as a colour-ordered n-gluon amplitude with $n-1$ on-shell momenta and one on-shell momentum. Recursion relations mitigate much of the computational overhead associated to tree-level matrix element calculations as any off-shell current in the process has to be evaluated only once for each phase space point.

An important property of tree-level matrix elements is the infrared behaviour. Due to the presence of propagators,

$$\mathcal{M}_n^{(0)}(\ldots,p_i,p_j,\ldots) \propto \frac{1}{(p_i+p_j)^2-m_{\widetilde{ij}}^2} \xrightarrow{\text{massless}} \frac{1}{2p_i\cdot p_j} = \frac{1}{2E_iE_j(1-\cos\theta_{ij})}\,, \tag{3.19}$$

amplitudes are divergent whenever an internal particles goes on-shell, $(p_i+p_j)^2 \to m_{\widetilde{ij}}^2$. This is the case when an emission becomes *unresolved*, *i.e.*, when the $n+1$-particle state becomes indistinguishable from the n-particle state. In the case where the emitted particle is massless, *cf.* eq. (3.19), this happens when the energy of the emission vanishes, the so-called *soft limit* $E_j \to 0$, or the two particles travel along the same axis, the so-called *collinear limit* $\theta_{ij} \to 0$. In these limits, (colour-ordered) amplitudes *factorise* into a reduced amplitude and universal collinear and soft *splitting functions* [191–194],

$$\mathcal{A}_n^{(0)}(\ldots,p_i^{h_i},p_j^{h_j},\ldots) \xrightarrow{i\parallel j} \sum_{h=\pm} \text{Split}_{-h}^{(0)}(p_i,p_j)\mathcal{A}_{n-1}^{(0)}(\ldots,p_{ij}^{h},\ldots)\,, \tag{3.20}$$

$$\mathcal{A}_n^{(0)}(\ldots,p_i,p_j,p_k,\ldots) \xrightarrow{p_j\to 0} \text{Soft}^{(0)}(p_i,p_j,p_k)\mathcal{A}_{n-1}^{(0)}(\ldots,p_i,p_k,\ldots)\,. \tag{3.21}$$

A similar factorisation holds also in the *quasi-collinear limit*, in which two massive partons travel along the same axis with energies much larger than their masses. The squared collinear and soft splitting functions build the basis for NLO subtraction schemes and parton showers and will therefore be discussed in more detail in section 3.1.4 and section 3.2. Similar factorisation properties hold in *double-unresolved limits*, in which two particles become simultaneously unresolved [195–199]. These become relevant for NNLO calculations and subleading-logarithmic corrections to parton showers.

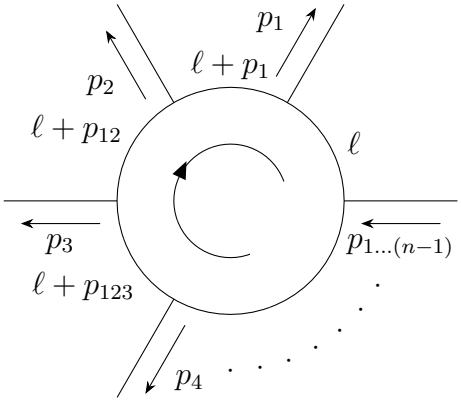


Figure 3.2: *Diagrammatic representation of a generic n-point one-loop integral. Arrows indicate the momentum flow through the diagram. Inspired by* [117, 200].

3.1.3 Loop Matrix Elements

Starting from the first order in the coupling constant[4], diagrams with closed internal loops contribute to the perturbative series in eq. (3.1), *cf.* fig. 3.1.

Since momenta of internal lines in Feynman diagrams cannot be measured, those can take any numerical value, constrained only by four-momentum conservation. For loop amplitudes such as the generic one sketched in fig. 3.2, this implies an integration over every loop momentum ℓ, so that one-loop amplitudes $\mathcal{M}_n^{(1)}$ contain integrals of the form

$$\mathcal{I}_n^D \propto \int \frac{\mathrm{d}^D \ell}{(2\pi)^D} \frac{1}{(\ell - m_1^2)((\ell + p_1)^2 - m_2^2)((\ell + p_{12})^2 - m_3^2) \cdots ((\ell + p_{1\cdots(n-1)})^2 - m_n^2)} \,, \tag{3.22}$$

where the notation $p_{1\ldots m} = \sum_{i=1}^{m} p_i$ is used. Here, we have reproduced only the simplest case of a scalar loop integral, while in general higher-rank integrals in which also the numerator depends on the loop momentum are required as well. The integral is written in $D \leq 4$ space-time dimensions as it is ill-defined in 4 dimensions, because the loop momentum ℓ can have arbitrary values and in particular may become infinitely large or vanish.

Within dimensional regularisation with $D = 4 - 2\varepsilon$ [201], the former leads to explicit poles of the form $1/\varepsilon$. These *ultraviolet* (UV) divergences can be removed by suitable redefinitions of all fields and couplings appearing in the Lagrangian, a method referred to as renormalisation, which introduces an arbitrary scale μ_R. The invariance of any physical observable on the choice of this scale gives rise to the renormalisation group equation eq. (2.25). Since for practical purposes, the perturbative series is truncated at finite order, residual dependencies on the renormalisation scale choice remain. It is therefore customary to estimate the size of these effects by scale variations of the form $\mu_\mathrm{R} \to k_\mathrm{R} \mu_\mathrm{R}$, where typically $k_\mathrm{R} \in \{1/2, 2\}$.

The latter type of divergences, appearing in the *infrared* limit $\ell^2 \to 0$, cannot be removed by such procedures. Instead, these are intrinsic to loop amplitudes and take universal forms in terms of explicit *soft* poles in $1/\varepsilon^2$ and *collinear* poles in $1/\varepsilon$. Following the notation introduced in [192], these universal infrared poles can be characterised as

[4] There exist so-called loop-induced processes, which have a loop already in the "Born" process. An example of such a process is the decay of a Higgs boson to two gluons, $\mathrm{H}^0 \to \mathrm{gg}$, which proceeds via an internal top-quark loop. Such amplitudes are, however, typically finite.

[192, 202, 203]

$$\mathcal{M}_n^{(1)}(\Phi_n;\mu_\mathrm{R},\mu_\mathrm{F}) = \mathbf{I}^{(1)}(\Phi_n,\mu_\mathrm{R},\varepsilon)\mathcal{M}_n^{(0)} + \mathcal{M}_n^{(1),\mathrm{finite}}\,, \tag{3.23}$$

with the one-loop singularity operator

$$\mathbf{I}^{(1)}(\Phi_n,\mu_\mathrm{R},\varepsilon) = \frac{1}{2}\frac{\mathrm{e}^{\varepsilon\gamma_\mathrm{E}}}{\Gamma(1-\varepsilon)}\sum_i \frac{1}{\mathbf{T}_i^2}\left(\mathbf{T}_i^2\frac{1}{\varepsilon^2}+\gamma_i\frac{1}{\varepsilon}\right)\sum_{j\neq i}\mathbf{T}_i\cdot\mathbf{T}_j\left(\frac{-s_{ij}}{\mu^2}\right)^{-\varepsilon}. \tag{3.24}$$

Here, $\mathbf{T}_i$ denotes the colour charge associated with the gluon emission off leg i and the anomalous dimensions γ_i are given by

$$\gamma_\mathrm{g} = \frac{11}{6}C_\mathrm{A} - \frac{2}{3}T_\mathrm{R}n_\mathrm{F}\,, \quad \gamma_\mathrm{q} = \gamma_{\bar{\mathrm{q}}} = \frac{3}{2}C_\mathrm{F}\,. \tag{3.25}$$

A similar characterisation formula exists for two-loop amplitudes [204].

The calculation of (the finite part of) one-loop matrix elements has been automated in publicly-available tools like OpenLoops [205], Recola [131,133], and MadLoop [206], using integral libraries [207–209] or integrand reduction tools [210–213]. These can be interfaced to event generators via the Binoth Les Houches Accord (BLHA) [214, 215], which defines a standard for the exchange of information between one-loop providers and event-generation frameworks. The central objective of such interfaces is to calculate the Born-one-loop interference in terms of a Laurent series expansion,

$$\begin{aligned}\mathrm{V}_n(\Phi_n,\mu_\mathrm{R},\mu_\mathrm{F}) &\equiv 2\,\mathrm{Re}\left[\mathcal{M}_n^{(0)}\mathcal{M}_n^{*(1)}\right](\Phi_n,\mu_\mathrm{R},\mu_\mathrm{F})\\ &= \frac{V_{-2}(\Phi_n,\mu_\mathrm{R},\mu_\mathrm{F})}{\varepsilon^2} + \frac{V_{-1}(\Phi_n,\mu_\mathrm{R},\mu_\mathrm{F})}{\varepsilon} + V_0(\Phi_n,\mu_\mathrm{R},\mu_\mathrm{F}) + \mathcal{O}(\varepsilon)\end{aligned} \tag{3.26}$$

for phase-space points $\Phi_n = \{p_1,\ldots,p_n\}$ defined by the event generator. A new BLHA interface to the extensive library of analytic one-loop amplitudes in MCFM will be presented in chapter 6.

As is the case for tree-level amplitudes, one-loop amplitudes factorise when emissions become unresolved. In the single-unresolved limits, the equivalent of eqs. (3.20) and (3.21) reads [198, 216–218]

$$\begin{aligned}\mathcal{A}_n^{(1)}(\ldots,p_i^{h_i},p_j^{h_j},\ldots) \xrightarrow{i\|j} \sum_{h=\pm}\Big(&\mathrm{Split}_{-h}^{(1)}(p_i,p_j)\mathcal{A}_{n-1}^{(0)}(\ldots,p_{ij}^{h},\ldots)\\ &+\mathrm{Split}_{-h}^{(0)}(p_i,p_j)\mathcal{A}_{n-1}^{(1)}(\ldots,p_{ij}^{h},\ldots)\Big)\,,\end{aligned} \tag{3.27}$$

$$\begin{aligned}\mathcal{A}_n^{(1)}(\ldots,p_i,p_j,p_k,\ldots) \xrightarrow{p_j\to 0} &\mathrm{Soft}^{(1)}(p_i,p_j,p_k)\mathcal{A}_{n-1}^{(0)}(\ldots,p_i,p_k,\ldots)\\ &+\mathrm{Soft}^{(0)}(p_i,p_j,p_k)\mathcal{A}_{n-1}^{(1)}(\ldots,p_i,p_k,\ldots)\,,\end{aligned} \tag{3.28}$$

with the tree-level splitting functions $\mathrm{Split}^{(0)}$, $\mathrm{Soft}^{(0)}$ given in eqs. (3.20) and (3.21) and new one-loop splitting functions $\mathrm{Split}^{(1)}$, $\mathrm{Soft}^{(1)}$. Here, $\mathcal{A}_n^{(0)}$ and $\mathcal{A}_n^{(1)}$ again denote colour-ordered amplitudes. Note that for one-loop ampitudes, "disconnected" colour topologies contribute, which, for pure-gluon amplitudes in the trace basis, are of the form [219]

$$\mathrm{Tr}(T^{a_1}\ldots T^{a_{j-1}})\,\mathrm{Tr}(T^{a_j}\ldots T^{a_n})\,. \tag{3.29}$$

The "squared" products of tree-level and one-loop splitting functions become relevant in the context of real-virtual corrections in NNLO calculations as well as virtual corrections to parton-shower branching kernels.

3.1.4 Subtraction Schemes

As stated in the previous two subsections, real-radiation and virtual matrix elements are separately divergent in four dimensions, with the latter exhibiting explicit poles in the dimensional regularisation parameter ε. A priori, it is thus not possible to perform NLO or higher-order calculations numerically in four dimensions. The Bloch-Nordsieck [220] and Kinoshita-Lee-Nauenberg [221, 222] theorems imply, however, that these divergences cancel between real and virtual corrections for *infrared-safe* observables. The numerical evaluation may thus be considered rather a technical difficulty than a prohibiting factor. The difficulty arises, because the virtual and real corrections are evaluated on different phase spaces: the former on the Born phase space, the latter on the real-radiation phase space.

Two distinct approaches have been derived to overcome this problem: phase-space slicing [202, 223] and subtraction [191, 192, 224–226]. While the former avoids dangerous regions in phase space by explicitly introducing an infrared cutoff (*i.e.*, "slicing" the phase space into resolved and unresolved regions), the latter explicitly subtracts singularities by utilising that amplitudes factorise in singular limits, *cf.* eqs. (3.20) and (3.21). Schematically, the subtracted $\mathcal{O}(\alpha_S)$ corrections to a differential Born cross section B_n read

$$\int \mathrm{d}\Phi_B \left[V_n(\Phi_B; \mu_R, \mu_F) + I_n(\Phi_B, \mu_R, \mu_F) \right]_{\varepsilon=0} + \int \mathrm{d}\Phi_R \left[R_n(\Phi_R; \mu_R, \mu_F) - S_n(\Phi_R; \mu_R, \mu_F) \right], \quad (3.30)$$

where I_n denotes the (analytical) integral of the subtraction term S_n over the factorised real-radiation phase space, schematically,

$$I_n(\Phi_B; \mu_R, \mu_F) = \int \mathrm{d}\Phi_{+1}\, S_n(\Phi_R(\Phi_B, \Phi_{+1}); \mu_R, \mu_F)\,, \quad (3.31)$$

Here, $\mathrm{d}\Phi_{+1}$ denotes the real-radiation phase-space measure which enters the factorised phase space,

$$\mathrm{d}\Phi_R = \mathrm{d}\Phi_B \times \mathrm{d}\Phi_{+1}\,. \quad (3.32)$$

The precise form of the mapping from the real-radiation phase space Φ_R to the Born phase space Φ_B depends on the subtraction scheme. Upon subtraction, the virtual and real correction terms in eq. (3.30) are separately finite in $D = 4$ dimensions.

The two most widely-used NLO subtraction schemes are the Frixione-Kunszt-Signer scheme [224, 225], which is automated in MadGraph_aMC@NLO [227], Whizard [228], and PowhegBox [229, 230]; and the Catani-Seymour scheme [191, 192, 231], which has been automated in MadGraph_aMC@NLO [232, 233], Sherpa [234], as well as in other packages [235–237]. Both schemes have been extended to treat NLO EW corrections [238, 239] and processes with coloured resonances [230, 240, 241]. In either method, the subtraction term is constructed based on the factorisation of squared matrix elements in single-unresolved limits, which follows directly from the amplitude-level factorisation eqs. (3.20) and (3.21).

As a result of eq. (3.20), a squared tree-level matrix element behaves as

$$\left|\mathcal{M}_n^{(0)}(\ldots,p_i,p_j,\ldots)\right| \xrightarrow{i\|j} \frac{8\pi\alpha_\mathrm{S}}{(p_i+p_j)^2-m_{\widetilde{ij}}^2}\left\langle \mathcal{M}_{n-1}^{(0)}(\ldots,p_{\widetilde{ij}},\ldots)\right| \hat{P}_{ij}(z) \left|\mathcal{M}_{n-1}^{(0)}(\ldots,p_{\widetilde{ij}},\ldots)\right\rangle \tag{3.33}$$

in the (quasi-)collinear limit. The scalar product is to emphasise spin correlations for gluon splittings $\mathrm{g}\mapsto\mathrm{gg}$ and $\mathrm{g}\mapsto\mathrm{q\bar{q}}$ and the functions $\hat{P}_{ij}$ denote the spin-dependent DGLAP splitting kernels, *cf. e.g.* [191]. It should be noted that the mass terms shield the singularity in the quasi-collinear limit. In the soft limit, it follows from eq. (3.21) that squared matrix elements behave like

$$\left|\mathcal{M}_n^{(0)}(\ldots,p_i,p_j,p_k,\ldots)\right| \xrightarrow{p_j\to 0} -4\pi\alpha_\mathrm{S}\sum_{i<k} S_{ik}(j)\left\langle \mathcal{M}_{n-1}^{(0)}(\ldots,p_i,p_k,\ldots)\right| \mathbf{T}_i\mathbf{T}_k \left|\mathcal{M}_{n-1}^{(0)}(\ldots,p_i,p_k,\ldots)\right\rangle , \tag{3.34}$$

where the non-abelian structure of QCD introduces non-trivial colour correlations via the insertion operator $\mathbf{T}_i\mathbf{T}_k$, *cf.* also eq. (3.24). We have here introduced the *eikonal* factor, which, in the massive case, is given by

$$S_{ik}(j)=\frac{2s_{ik}}{s_{ij}s_{jk}}-\frac{2m_i^2}{s_{ij}^2}-\frac{2m_k^2}{s_{jk}^2}=\frac{2(p_i\cdot p_k)}{2(p_i\cdot p_j)2(p_j\cdot p_k)}-\frac{2p_i^2}{(2p_i\cdot p_j)^2}-\frac{2p_k^2}{(2p_j\cdot p_k)^2}\,. \tag{3.35}$$

In the Catani-Seymour scheme, the soft eikonal factor is partitioned so that *dipole functions* $V_{ij,k}$ can be defined, accounting for the $i-j$ (soft-)collinear limit. Together with a suitable *recoil scheme*, *i.e.*, a prescription for how to recombine momenta $p_i+p_j+p_k=p_{\widetilde{ij}}+p_{\widetilde{k}}$, Catani-Seymour dipole subtraction terms are constructed as [191, 192, 231]

$$S_n^{ij,k}=-\frac{8\pi\alpha_\mathrm{S}}{(p_i+p_j)^2-m_{\widetilde{ij}}^2}\left\langle \mathcal{M}_{n-1}^{(0)}(\ldots,p_{\widetilde{ij}},p_{\widetilde{k}},\ldots)\right| \frac{\mathbf{T}_{ij}\mathbf{T}_k}{\mathbf{T}_{ij}^2}V_{ij,k}(p_i,p_j,p_k) \left|\mathcal{M}_{n-1}^{(0)}(\ldots,p_{\widetilde{ij}},p_{\widetilde{k}},\ldots)\right\rangle , \tag{3.36}$$

The full subtraction term S_n is then constructed as a sum of $S_n^{ij,k}$ over all dipoles and I_n as a sum of dipole functions, analytically integrated over their respective dipole phase space in $D=4-2\varepsilon$ dimensions,

$$\mathcal{V}_{ij,k}(\varepsilon)=\int \mathrm{d}^{4-2\varepsilon}\Phi_{+1}^{ij,k}(p_i,p_j,p_k)\,V_{ij,k}(p_i,p_j,p_k) \tag{3.37}$$

where we have used the exact phase-space factorisation

$$\mathrm{d}^{4-2\varepsilon}\Phi_n = \mathrm{d}^{4-2\varepsilon}\Phi_{+1}^{ij,k}\,\mathrm{d}\Phi_{n-1}\,. \tag{3.38}$$

A set of spin-averaged final-final dipole functions $V_{ij,k}$ will be given in section 3.2.2.

A particularly appealing subtraction scheme is given by the antenna formalism [226], as the NLO scheme [242–244] naturally extends to the NNLO [245–250]. The basic idea is to use colour-ordered decay matrix elements as subtraction terms, since these naturally include all single- and double-unresolved limits. The NLO *antenna functions* are thus

defined by [245]

$$X_3^0(p_i,p_j,p_k) = \frac{\left|\mathcal{A}_3^{(0)}(p_i,p_j,p_k)\right|^2}{g_S^2\left|\mathcal{A}_2^{(0)}(p_I,p_K)\right|^2}\,. \tag{3.39}$$

As opposed to the dipole functions described above, antenna subtraction starts from the soft limit and fractions the collinear limit into neighbouring antennae. As such, the antenna functions defined by eq. (3.39) are typically decomposed into a linear combination of *sub-antenna functions* (*global* antenna functions), each of which contains the full soft, but only part of the collinear limit in a given antenna $i-j-k$. Equivalent to the case for dipole subtraction, the full subtraction term S_n is constructed as a sum of antenna subtraction terms,

$$S_n^{j/IK} = 4\pi\alpha_S \mathcal{C}_{j/IK} X_3^0(p_i,p_j,p_k)\left|\mathcal{M}_{n-1}^{(0)}(\ldots,p_I,p_K,\ldots)\right|, \tag{3.40}$$

where $\mathcal{C}_{j/IK}$ denotes the colour factor of the branching $IK \mapsto ijk$. The virtual subtraction term I_n is again constructed as the sum over integrated antenna functions,

$$\mathcal{X}_3^0(\varepsilon) = \int \mathrm{d}^{4-2\varepsilon}\Phi_{+1}^{j/IK}(p_i,p_j,p_k)\, X_3^0(p_i,p_j,p_k)\,, \tag{3.41}$$

with the $(4-2\varepsilon)$-dimensional antenna phase space $\mathrm{d}^{4-2\varepsilon}\Phi_{+1}^{j/IK}$, for which the exact factorisation

$$\mathrm{d}^{4-2\varepsilon}\Phi_n = \mathrm{d}^{4-2\varepsilon}\Phi_{+1}^{j/IK}\,\mathrm{d}\Phi_{n-1}\,, \tag{3.42}$$

holds. Similarly, at NNLO, the double-real and real-virtual antenna functions are defined as [245]

$$X_4^0(p_i,p_j,p_k,p_\ell) = \frac{\left|\mathcal{A}_4^{(0)}(p_i,p_j,p_k,p_\ell)\right|^2}{\left|\mathcal{A}_2^{(0)}(p_I,p_K)\right|^2}\,, \tag{3.43}$$

$$X_3^1(p_i,p_j,p_k) = \frac{\left|\mathcal{A}_3^{(1)}(p_i,p_j,p_k)\right|^2}{\left|\mathcal{A}_2^{(0)}(p_I,p_K)\right|^2} - X_3^0(p_i,p_j,p_k)\frac{\left|\mathcal{A}_2^{(1)}(p_I,p_K)\right|^2}{\left|\mathcal{A}_2^{(0)}(p_I,p_K)\right|^2} \tag{3.44}$$

Antenna functions for quark-antiquark parents can be calculated from off-shell photon decays $\gamma^* \to \mathrm{q\bar{q}}$ [251–254], quark-gluon antennae from neutralino decays to a gluino[5] and a gluon $\tilde{\chi} \to \tilde{\mathrm{g}}\mathrm{g}$ [255], and gluon-gluon antennae from Higgs decays to two gluons in the Higgs effective theory[6] $\mathrm{H}^0 \to \mathrm{gg}$ [256]. Antenna functions for configurations containing initial-state partons can be derived from these by crossing invariance, *i.e.*, by crossing final-state legs into the initial state, flipping their momenta $p_i \to -p_i$. An extensive review of the full subtraction formalism up to the NNLO can be found in [257]. Shower models based on the antenna formalism will be discussed in section 3.2.3.

[5] In supersymmetric extensions of the Standard Model, the gluino is the fermionic partner of the gluon. In any collinear limit it thus behaves like a quark, although it transforms as an octet under the colour group and as such contains a colour and an anticolour index in the colour-flow basis. This fact leads to spurious singularities in the quark-gluon antenna functions which have to be accounted for when using them as subtraction terms.

[6] In the Higgs effective theory (HEFT), the top-quark loop is integrated out, leaving an effective $\mathrm{H}^0 - \mathrm{g} - \mathrm{g}$ vertex.

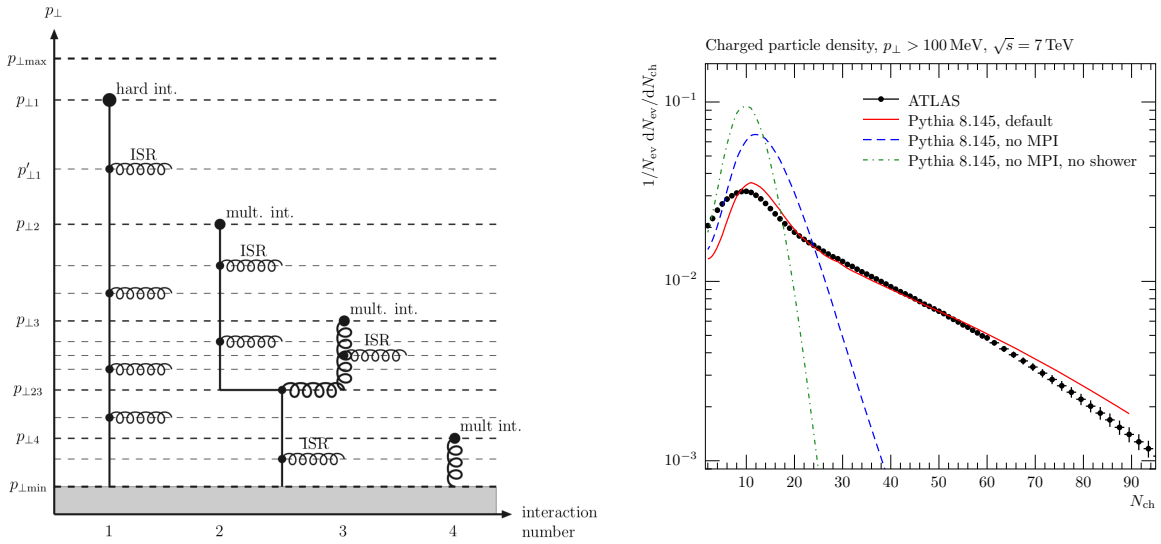


Figure 3.3: *Illustration of MPIs interleaved with the evolution of the primary and further hard processes in a common measure of transverse momentum (left). Importance of MPIs in the description of the charged particle density compared to the distribution measured in the ATLAS experiment* [264]. *Taken from* [262] *and* [90], *respectively.*

3.1.5 Multiple Hard Interactions

When thinking of hadrons as containers for quarks and gluons, it might seem counter-intuitive to restrict a hadron-hadron collision only to a single parton-parton collision process. In fact, for some observables the effect of additional (semi-) hard processes can be significant, as shown in the right-hand pane of fig. 3.3. Multi-purpose event generators therefore usually include the modelling of multi-parton interactions (MPIs).

The first model to take into account the composite nature of protons was developed in [258] and assumed fairly independent interactions between the proton composita. Nowadays, MPI models in PYTHIA and SHERPA are based on p_{T}-ordered Sudakov evolution [259, 260] based on the observation that the cross section of pure-QCD processes behaves as

$$\mathrm{d}\sigma_{\text{QCD}} \propto \frac{\mathrm{d}p_{\text{T}}^2}{p_{\text{T}}^4}\,, \tag{3.45}$$

whereas the model implemented in HERWIG is based on eikonal factorisation [261].

In PYTHIA, the generation of MPIs is interleaved with the parton showers [262, 263], *i.e.*, both initial- and final-state showers and MPIs compete for the highest scale and any subsequent hard process is colour-connected to the primary one and its shower emissions. This procedure is schematically depicted in fig. 3.3. The models in SHERPA and HERWIG are independent of the rest of the event simulation. In all of the three generators, MPIs are subject to parton-shower radiation.

While all of the above-mentioned MPI models are based on judicious assumptions, they are not derived from first principles, as is the case for (the single-parton interaction) fixed-order calculations discussed previously in this section. A rigorous approach to double-parton scattering as the simplest MPI has been developed in [265–269] and recently implemented in a Monte Carlo program, including an NLO subtraction scheme which is matched to a parton shower [270, 271].

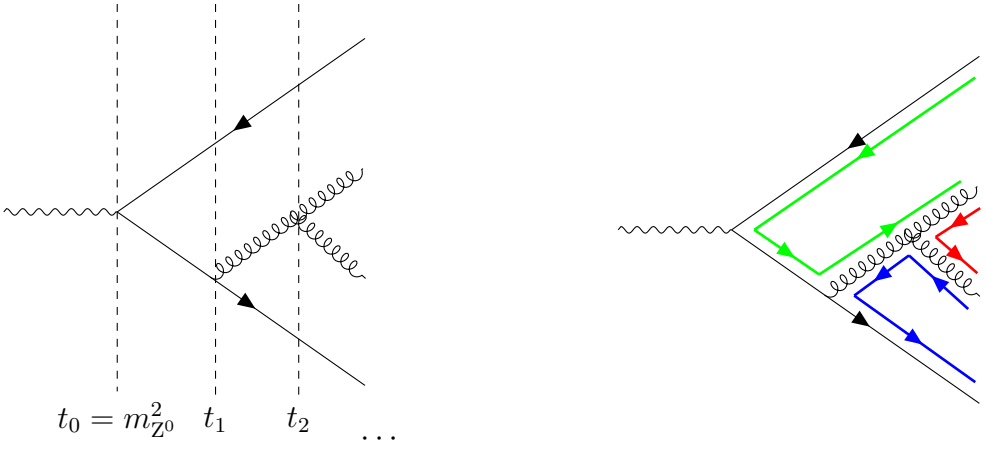


Figure 3.4: *Illustration of a parton shower in hadronic* Z^0 *decays* $Z^0 \to q\bar{q}$ *(left) and associated colour flow in the large-*N_C *limit (right).*

3.2 Parton Showers and Resummation

In hard scattering processes described in the previous subsection, a handful of partons are created at high energies and with large momenta. The detector signal, on the other hand, contains a large number of particles. The transition from the former to the latter occurs via successive *bremsstrahlung*, a process in which the initially hard partons lose energy through *parton branchings*. According to the Feynman rules of QCD, *cf.* section 2.2.1, three such branchings may occur for a final-state parton[7]: $g \mapsto gg$, $g \mapsto q\bar{q}$, and $q \mapsto qg$. In the latter, one may be distinguish between the branching $q \mapsto qg$ in which the gluon can become soft and $q \mapsto gq$, where the gluon is considered the hard radiator. Similar $1 \mapsto 2$ branchings can be derived for the QED/EW case. In Monte Carlo event generators, this evolution between the high- and low-energy regime is described by a *parton shower*. An illustration of a parton shower describing *final-state radiation* (FSR) in the decay $Z^0 \to q\bar{q}$ is shown in fig. 3.4. If the beam particles are hadrons or QED/EW evolution is considered, particle branchings may also occur in the initial state, a process called *initial-state radiation* (ISR). In parton showers, ISR is typically modelled as a conditional backwards evolution [272], *i.e.*, starting with the hard process and evolving the initial-state legs backwards in time under the condition that any occurring branching must not change the original initial-state parton configuration (flavour, momentum, colour, and spin).

Parton-shower algorithms build upon the factorisation of squared matrix elements[8] in the soft and collinear limits, *cf.* eqs. (3.33) and (3.34),

$$\mathrm{d}\sigma_{n+1} \xrightarrow{\text{single-unresolved}} \sum_j K_{j/IK}\, \mathrm{d}\Phi^j_{+1}\, \mathrm{d}\sigma_n\,, \tag{3.46}$$

where $K_{j/IK}$ denotes a generic branching kernel

$$K_{j/IK}(t,\zeta,\phi) = 4\pi\alpha_S(t)\, R_{\mathrm{PDF}}(t,\zeta)\, \mathcal{C}_{j/IK}\, \bar{K}_{j/IK}(t,\zeta,\phi) \tag{3.47}$$

for the branching $IK \mapsto ijk$ and $\bar{K}_{j/IK}$ its colour- and coupling-stripped equivalent. We have here introduced *shower variables* $\{t,\zeta,\phi\}$, in terms of which we can write the radiation phase space $\mathrm{d}\Phi_{+1}$ in the factorisation of the n-parton phase-space, *cf.* eq. (3.32),

$$\mathrm{d}\Phi_n(p_a,p_b;p_1,\ldots,p_n) = \mathrm{d}\Phi^j_{+1}(p_i,p_j,p_k)\, \mathrm{d}\Phi_{n-1}(p_a,p_b;\tilde{p}_1,\ldots,\tilde{p}_n)\,, \tag{3.48}$$

[7] We will denote branchings by $a \mapsto bc$ to distinguish them from reactions $a \to bc$.

[8] Recently, parton evolution on the amplitude level has been considered, *cf. e.g.* [273, 274].

$$\mathrm{d}\Phi^{j}_{+1}(p_i,p_j,p_k) \equiv \mathrm{d}\Phi^{j}_{+1}(t,\zeta,\phi) = \frac{1}{16\pi^2}|J(t,\zeta,\phi)|\,\mathrm{d}t\,\mathrm{d}\zeta\,\mathrm{d}\phi\,. \tag{3.49}$$

In this context, the variable t denotes some "singularity measure", *i.e.*, a measure of softness or collinearity, typically defined such that $t \to 0$ in the unresolved limits. It follows that the probability $\mathcal{P}(t_0,t)$ for a branching to occur at any scale t' with $t_0 \geq t' > t$ can be written as

$$\mathcal{P}(t_0,t) = \sum_j \int_t^{t_0} \int_{\zeta_{\min}}^{\zeta_{\max}} \int_0^{2\pi} \mathrm{d}t'\,\mathrm{d}\zeta'\,\mathrm{d}\phi'\,|J(t',\zeta',\phi')|K_{j/IK}(t',\zeta',\phi')\,. \tag{3.50}$$

Naively, eq. (3.50) could be applied iteratively to approximate the cross section of higher-multiplicity states. Such a procedure would, however, lead to a diverging inclusive cross section and is therefore not useful. This can be circumvented by imposing *unitarity* on the shower evolution, made explicit in terms of a *no-branching* probability, often called Sudakov (form) factor, $\Pi_{j/IK}(t_0,t)$, so that

$$\mathcal{P}_{j/IK}(t_0,t) + \Pi_{j/IK}(t_0,t) = 1\,. \tag{3.51}$$

Assuming a Markovian branching process, *i.e.*, that successive branchings have no memory of the previous ones, the no-branching probability between the shower starting scale t_0 and the branching scale t can be written as the limit of a product of no-branching probabilities over decreasingly small evolution windows $[\frac{m(t_0-t)}{N}, \frac{(m+1)(t_0-t)}{N}]$,

$$\Pi_{j/IK}(t_0,t) = \lim_{N\to\infty} \prod_{m=0}^{N-1} (1 - \mathcal{P}(t_m,t_{m+1})) = \exp\left\{ -\int_t^{t_0} \mathrm{d}t'\,\frac{\mathrm{d}\mathcal{P}(t')}{\mathrm{d}t'} \right\}. \tag{3.52}$$

As we shall see below, the branching kernels $K_{j/IK}$ involve terms inversely proportional to t, so that the term in the exponential above is of the form

$$\frac{\mathrm{d}\mathcal{P}(t)}{\mathrm{d}t} \propto \frac{\mathrm{d}t}{t}\,, \tag{3.53}$$

where a potentially complicated function of t, ζ, and ϕ multiplies the right-hand side. Equation (3.52) therefore *resums*[9] logarithmic terms of the form $\log(t_0,t)$, which become large if $t_0 \gg t$. Moreover, considering a gluon emission $\mathrm{q} \mapsto \mathrm{qg}$ or $\mathrm{g} \mapsto \mathrm{gg}$, in the limit of the emission being collinear and soft, the splitting kernels behave like $1/(1-\zeta)$, assuming that the splitting variable ζ denotes a measure of the momentum fraction of the emission in this limit. The integration over ζ therefore introduces a further logarithm, so that the *leading logarithms* in this limit are given by a term of the form $\alpha_{\mathrm{S}}\log^2(t_0/t)$.

Generalising eq. (3.52) to the evolution of an n-parton state at scale t_n to an $(n+1)$-parton state at scale t_{n+1} yields

$$\begin{aligned}\Pi_n(t_n,t_{n+1}) &= \exp\left\{ -\sum_j \int_{t_{n+1}}^{t_n} \int_{\zeta_{\min}}^{\zeta_{\max}} \int_0^{2\pi} \mathrm{d}t'\,\mathrm{d}\zeta'\frac{\mathrm{d}\phi'}{2\pi}\,\frac{1}{16\pi^2}|J(t,\zeta,\phi)|K_{j/IK}(t',\zeta',\phi') \right\} \\ &= \prod_j \exp\left\{ -\int_{t_{n+1}}^{t_n} \int_{\zeta_{\min}}^{\zeta_{\max}} \int_0^{2\pi} \mathrm{d}t'\,\mathrm{d}\zeta'\frac{\mathrm{d}\phi'}{2\pi}\,\frac{1}{16\pi^2}|J(t,\zeta,\phi)|K_{j/IK}(t',\zeta',\phi') \right\}\end{aligned} \tag{3.54}$$

[9] Resummation here refers to reorganising, *i.e.*, *re-summing*, the perturbative series in α_{S}^n into terms of the form $\alpha_{\mathrm{S}}^n \log^m(t_0/t)$.

$$= \prod_j \Pi_{j/IK}(t_n, t_{n+1}) \,,$$

for which we introduce the shorthand

$$\Pi_n(t_n, t_{n+1}) = \exp\left\{ - \int_{t_{n+1}}^{t_n} \mathrm{d}\Phi_{+1}(t', \zeta', \phi') \, \mathrm{K}_{n\mapsto n+1}(\Phi_{+1}(t', \zeta', \phi')) \right\} . \tag{3.55}$$

To describe successive emissions by a (unitary) parton shower, we introduce a *generating functional* $\mathcal{S}_n$, the so-called *shower operator*, which is recursively defined by

$$\mathcal{S}_n(t, O) = \Pi_n(t, t_\mathrm{c}) O(\Phi_n) + \int_{t_\mathrm{c}}^{t} \mathrm{d}\Phi_{+1}(t', \zeta', \phi') \, \mathrm{K}_{n\mapsto n+1} \Pi_n(t, t') \mathcal{S}_{n+1}(t', O) \,. \tag{3.56}$$

We have here introduced a *shower cut-off* scale $t_\mathrm{c} \approx \Lambda^2_\mathrm{QCD} \approx 1\ \mathrm{GeV}^2$ to ensure that partonic processes are calculated only in the regime of perturbative QCD[10] and to make the parton-shower algorithm amenable to Monte Carlo methods. Below the cut-off scale singularities of unresolved real contributions cancel with the ones of virtual corrections, *cf.* section 3.1.4. Above the cut-off scale, real corrections are explicitly modelled by the parton shower and virtual corrections are implicitly modelled by assuming that they exactly cancel the real contributions. In addition, the shower evolution defined by eq. (3.56) is required to be *strongly ordered*, *i.e.*, successive branching scales have to be decreasing, $t_0 > t_1 > t_2 > \cdots > t_\mathrm{c}$, *cf.* fig. 3.4. In this way, the shower algorithm accounts for the dominant contributions of any given amplitude, namely the ones in which propagators of successive branchings become more and more on-shell. These also correspond to the leading single-unresolved singularities. First approaches to include unordered contributions in parton showers have been developed based on direct double-branchings [275, 276]. These become important in the context of matching parton showers to NNLO calculations, as will be discussed in chapter 7.

Adding the effect of a parton shower to the LO calculation of an infrared-safe observable O, *cf.* eq. (2.57), therefore gives

$$\langle O \rangle_n^\mathrm{PS} = \int \mathrm{d}\Phi_\mathrm{B} \, \mathrm{B}_n(\Phi_\mathrm{B}; \mu_\mathrm{R}, \mu_\mathrm{F}) \mathcal{S}_n(t, O) \,. \tag{3.57}$$

That the shower operator, eq. (3.56), is manifestly unitary can be seen as follows. Ignoring for the time being the recursive appearance of the shower operator, the first term in eq. (3.56) describes the probability that no branching happens between the scales t and the shower cutoff t_c, while the second term in eq. (3.56) describes the probability that a branching occurs at any scale $t' \in [t, t_\mathrm{c}]$. As such, eq. (3.56) is equivalent to eq. (3.51). To prove that this holds to all orders, this argument can be repeated by induction. In practice, unitarity is implemented in shower Monte Carlos by explicitly evolving an n-parton state to an $(n+1)$-parton state upon a branching. This means, that whenever a parton in the n-parton state branches, this state is "destroyed" and replaced by the $(n+1)$-parton state.

Commonly, parton showers are formulated in a large-N_C limit, often referred to as the *leading colour* limit, *cf.* the right hand pane in fig. 3.4. This means that, before the shower evolution commences, the colour structure of the hard process is projected

[10] This does not apply *e.g.* to QED showers, which can be continued below the hadronisation scale. There, a cut-off scale is needed to ensure termination of the algorithm and to avoid proliferation of unresolvable emissions.

onto a planar colour flow (*cf.* section 3.1.2 for a discussion of the colour flow basis), *i.e.*, one which can be drawn in 2D without crossing colour lines. In the shower, only emissions off neighbouring colour-anticolour pairs (*colour dipoles*) are considered. While empirically this procedure has been shown to give reasonable predictions in many cases, it systematically misses some contributions which are suppressed by $1/N_C^2$ relative to the leading-colour factor. In terms of the colour matrix $\mathbf{T}_i\mathbf{T}_k$ introduced in section 3.1.3 and section 3.1.4, this means that the shower only considers branchings happening along its diagonal. To finite order, the effect of sub-leading colour correlations can be taken into account by matrix-element corrections [277–282] (*cf.* section 3.3.1 for a discussion of matrix-element corrections), *i.e.*, by iteratively reweighting branchings to the correct colour structure. This does not generically account for colour correlations in the full shower, *i.e.*, to infinite order. Approaches addressing the issue to all orders have been presented in [283, 284], where the "most important" sub-leading colour corrections are systematically included, and [285], which is based on arguments stemming from soft-gluon resummation. Since colour correlation matrices $\mathbf{T}_i\mathbf{T}_k$ quickly become intractable, work has begun to formulate parton showers directly on the amplitude level [273, 274, 286]. In any of the approaches, hadronisation corrections (*cf.* section 3.4) require special attention, as current hadronisation models are intrinsically linked to the large-N_C limit and no generalisation with an exact colour treatment exists yet.

A lot of recent research has considered the *logarithmic accuracy* of parton showers [281, 282, 287–293]. As we have seen above, showers account for large logarithms in a reorganised (resummed) perturbative series. It is, however, usually not clear which terms exactly a certain shower algorithm resums in this series. To understand this statement, let us consider the cumulative distribution $\Sigma(L)$, *i.e.*, the probability for an observable to have a value smaller than $\exp(-L)$, where L denotes a logarithm that depends on the ratio of the hard scale t_0 and the resolved scale t. For a certain class of observables[11], $\Sigma(L)$ can be written as [288, 295]

$$\Sigma(L) = \exp\left\{Lg_1(\alpha_S L) + g_2(\alpha_S L) + \alpha_S g_3(\alpha_S L) + \ldots\right\} . \tag{3.58}$$

Here, the first term resums the *leading logarithms* (LL) of the form $\alpha_S^n L^{n+1}$, the second term the *next-to-leading logarithms* (NLL) $\alpha_S^n L^n$, etc. Returning to the question of the logarithmic accuracy of parton showers, it is often assumed that they generically perform an LL resummation, independently of the observable considered. Since it has been shown that this is not always true [281, 282, 288, 289, 291], the assessment of the shower accuracy has become an important aspect of precision physics.

In the following three subsections, common parton-shower implementations will briefly be discussed and compared. The discussion is organised in terms of the different factorisation paradigms employed in different showers, although this does not take the choice of ordering variables, recoil schemes, and the treatment of colour correlations into account. We will not describe parton showers derived from quantum density matrices, as implemented in the DEDUCTOR program [296]; an exhaustive review of that formalism can be found in [297].

3.2.1 DGLAP Evolution

Historically, parton showers were first constructed based on collinear factorisation. In traditional parton showers, the branching kernels are the DGLAP splitting kernels [73,

[11] Typically many so-called "event shapes" and some "jet rates", *cf. e.g.* [294] for more details.

86, 87], in the massless version given by

$$P_{\mathrm{qq}}(z) = C_{\mathrm{F}}\left[\frac{2}{1-z} - (1+z)\right], \tag{3.59}$$

$$P_{\mathrm{gg}}(z) = 2C_{\mathrm{A}}\left[\frac{1}{1-z} + \frac{1}{z} + z(1+z) - 2\right], \tag{3.60}$$

$$P_{\mathrm{gq}}(z) = T_{\mathrm{R}}\left[1 + 2z(1-z)\right], \tag{3.61}$$

where the notation P_{Ii} is used to denote the branching $I \mapsto ij$, *i.e.* the emitted particle is kept implicit. The evolution equations take the form

$$\mathrm{d}\mathcal{P}_{Ii}(t) = \frac{\alpha_{\mathrm{S}}(t)}{2\pi}\frac{\mathrm{d}t}{t}\int_{z_{\min}}^{z_{\max}} \mathrm{d}z P_{Ii}(z), \tag{3.62}$$

$$\mathrm{d}\mathcal{P}_{Ii}(x,t) = \frac{\alpha_{\mathrm{S}}(t)}{2\pi}\frac{\mathrm{d}t}{t}\int_{z_{\min}}^{z_{\max}} \frac{\mathrm{d}z}{z}\frac{f_{I/h}(x/z,t)}{f_{i/h}(x,t)}P_{Ii}(z), \tag{3.63}$$

where the evolution variable t may be chosen proportional to the virtuality of the branching particle, the transverse momentum of the emission, or a generalised (energy-weighted) emission angle. Equation (3.63) has already been given in the context of PDF evolution in eq. (2.56) in slightly different form. As such, DGLAP-based parton showers may be viewed as a numerical implementation of the parton evolution encoded in the DGLAP equation eq. (2.56). There is, however, an important subtlety: typically, parton showers require an on-shell evolution, meaning that at each step in the branching sequence, every parton is on its mass shell. This cannot be fulfilled by $1 \mapsto 2$ branchings unless the emission and the emitter are exactly collinear. Therefore, typically some momentum is taken from the colour-partner of the branching parton ("local recoil") or the whole event ("global recoil") in order to ensure momentum conservation in the branching process.

There is a related subtlety: because DGLAP showers derive from the collinear limit, they are a priori not accurate for soft, wide-angle emissions. On average, these effects can be taken into account by *angular ordering*, which implies that the effect of the soft eikonal, eq. (3.35), is limited to a cone of opening angle $\sim 2\theta$ around the emitting leg. In case a soft gluon is emitted off a collinear gluon-gluon pair, the emission interferes destructively outside this cone. A coherent emission pattern may therefore be restored in DGLAP showers either by ordering emissions in angle [298, 299] or by explicitly vetoing emissions at too large angles [300].

Parton showers based on DGLAP splitting kernels have previously been implemented in SHERPA's APACIC++ [301, 302] shower and still form the basis of HERWIG's angular ordered $\tilde{q}$ shower [303] and PYTHIA's $p_\perp$-ordered shower [262].

3.2.2 Dipole Showers

The idea of Catani-Seymour dipole showers, often simply called *dipole showers*, is to use the dipole subtraction terms introduced in section 3.1.4 as shower branching kernels. Because of this, dipole showers are particularly suited for matching showers to NLO calculations [304, 305]. This topic will be discussed in more detail in section 3.3.2.

For final-state branchings with a final-state spectator, the massless, spin-averaged

dipole functions[12] are given by

$$V_{\mathrm{q_i g_j},k}(y_{ij,k}, z_i) = C_\mathrm{F}\left[\frac{2}{1-z_i(1-y_{ij,k})} - (1+z_i)\right], \tag{3.64}$$

$$V_{\mathrm{g_i g_j},k}(y_{ij,k}, z_i) = 2C_\mathrm{A}\left[\frac{1}{1-z_i(1-y_{ij,k})} + \frac{1}{1-(1-z_i)(1-y_{ij,k})} + z_i(1+z_i) - 2\right], \tag{3.65}$$

$$V_{\mathrm{q_i g_j},k}(y_{ij,k}, z_i) = T_\mathrm{R}\left[1 - 2z_i(1-z_i)\right], \tag{3.66}$$

expressed in terms of the two invariants

$$y_{ij,k} = \frac{s_{ij}}{s_{ij}+s_{jk}+s_{ik}}, \quad z_i = \frac{s_{ik}}{s_{ik}+s_{jk}}. \tag{3.67}$$

The full set of massless and massive dipole functions can be found in [192] and [231], respectively. Dipole showers are typically evolved in a measure of transverse momentum $k_\perp$, which in the case of massless final-state branchings is given by

$$k_\perp^2 = y_{ij,k} z_i (1-z_i) Q^2, \quad Q^2 = (p_i + p_j + p_k)^2. \tag{3.68}$$

The kinematics are constructed as exact inverses of the dipole kinematics in the subtraction scheme, *cf.* section 3.1.4. For massless final-state branchings with final-state recoilers, the local dipole kinematics are constructed according to the following Sudakov parametrisation

$$p_i = z_i \tilde{p}_{ij} + \frac{k_\perp^2}{z_i Q^2}\tilde{p}_k + k_\perp, \tag{3.69}$$

$$p_j = (1-z_i)\tilde{p}_{ij} + \frac{k_\perp^2}{(1-z_i)Q^2}\tilde{p}_k - k_\perp, \tag{3.70}$$

$$p_k = (1-y_{ij,k})\tilde{p}_k. \tag{3.71}$$

As for DGLAP-based showers, *cf.* section 3.2.1, in the case of initial-state branchings with initial-state spectators, a global recoil scheme has to be employed to ensure that the initial-state partons remain aligned with the beam axis. Global recoil schemes are sometimes also employed in other kinematical setups [290, 291, 306, 307].

Dipole showers along the lines described above are implemented in SHERPA [308] and HERWIG [120, 309].

There is an alternative formulation of dipole showers, which is more akin to traditional parton showers, while borrowing concepts from antenna showers, *cf.* section 3.2.3. It has been derived in the context of the DIRE [13] shower [310], implemented as non-default shower options in PYTHIA 8.3 [14] as well as SHERPA. The massless final-final branching kernels in DIRE can be written analogously to eqs. (3.59) to (3.61),

$$P_\mathrm{qq}(z, \kappa^2) = C_\mathrm{F}\left[\frac{2(1-z)}{(1-z)^2+\kappa^2} - (1+z)\right], \tag{3.72}$$

$$P_\mathrm{qg}(z, \kappa^2) = C_\mathrm{F}\left[\frac{1+(1-z)^2}{z}\right], \tag{3.73}$$

[12] We systematically exclude coupling factors in branching kernels.

[13] Dipole-Resummation

[14] With the release of PYTHIA 8.3, DIRE for PYTHIA was incorporated into the main PYTHIA code, but was maintained as a plug-in to PYTHIA 8.2 before.

$$P_{\mathrm{gg}}(z,\kappa^2) = 2C_{\mathrm{A}}\left[\frac{1-z}{(1-z)^2+\kappa^2} + \frac{1}{z} + z(1+z) - 2\right], \tag{3.74}$$

$$P_{\mathrm{gq}}(z,\kappa^2) = T_{\mathrm{R}}\left[1 + 2z(1-z)\right], \tag{3.75}$$

with the shower ordering and splitting variables

$$\kappa^2 = \frac{s_{ij}s_{jk}}{Q^4}, \quad z = 1-(1-z_i)(1-y_{ij,k}) \tag{3.76}$$

where the former here has a similar effect as the regulator in principal-value regularisation [311]. Note that this definition is different to eq. (3.68), but identical to (a dimensionless version of) the evolution variable in antenna showers eq. (3.84). Due to the close connection with both DGLAP parton evolution and the dipole framework, it has been possible to implement NLO DGLAP [312], including triple-collinear splitting functions [276], and two-loop soft corrections [313] in DIRE.

3.2.3 Antenna Showers

Historically the first "dipole" shower was the Lund dipole model implemented in ARIADNE [314, 315]. The formalism, nowadays referred to as (dipole-) antenna showers, has been revived in the VINCIA [277,316,317] plug-in to PYTHIA and the Krauss-Winter shower [318] in an early version of SHERPA.

In contrast to DGLAP and dipole showers, *cf.* sections 3.2.1 and 3.2.2, antenna showers do not start from the factorisation in the collinear limit but from the soft factorisation eq. (3.34). As such, the starting point is the soft eikonal eq. (3.35) and antennae are intrinsically agnostic to the assignment of a recoiler and an emitter. Instead, both parent partons act simultaneously as both, the emitter and recoiler. While this treatment is adverse to the parton-evolution picture maintained in DGLAP and dipole showers, it allows for coherence to be incorporated into the evolution in a simple way. Moreover, since always a parton *pair* emits at a time, cross sections factorise into a smaller number of branching kernels/subtraction terms $K_{j/IK}$ in the antenna formalism.

The branching kernels, called *antenna functions* in this context, can be derived from the definition eq. (3.39). In terms of the two dimensionless invariants

$$y_{ij} = \frac{s_{ij}}{s_{ij}s_{jk}}, \quad y_{jk} = \frac{s_{jk}}{s_{ij}s_{jk}}, \tag{3.77}$$

and the antenna mass $s_{IK} = s_{ij} + s_{jk} + s_{ik}$, a set of massless final-final (FF) antenna functions is given by

$$\bar{A}_{\mathrm{g}_j/\mathrm{q}_I\bar{\mathrm{q}}_K}(y_{ij},y_{jk}) = \frac{1}{s_{IK}}\left[\frac{(1-y_{ij})^2+(1-y_{jk})^2}{y_{ij}y_{jk}} + 1\right], \tag{3.78}$$

$$\bar{A}_{\mathrm{g}_j/\mathrm{q}_I\mathrm{g}_K}(y_{ij},y_{jk}) = \frac{1}{s_{IK}}\left[\frac{(1-y_{ij})^3+(1-y_{jk})^2}{y_{ij}y_{jk}} + 2 - y_{ij} - \frac{y_{jk}}{2}\right], \tag{3.79}$$

$$\bar{A}_{\mathrm{g}_j/\mathrm{g}_I\mathrm{g}_K}(y_{ij},y_{jk}) = \frac{1}{s_{IK}}\left[\frac{(1-y_{ij})^3+(1-y_{jk})^3}{y_{ij}y_{jk}} + 3 - \frac{3y_{ij}}{2} - \frac{3y_{jk}}{2}\right], \tag{3.80}$$

$$\bar{A}_{\mathrm{q}_j/\mathrm{g}_I X_K}(y_{ij},y_{jk}) = \frac{1}{s_{IK}}\frac{1}{2y_{ij}}\left[y_{jk}^2 + (1-y_{ij}-y_{jk})^2\right], \tag{3.81}$$

where we have omitted colour factors. The assignment of these for gluon-emission antenna functions is less straight-forward, because they cannot be derived from the collinear limit (as for DGLAP and dipole kernels) and the soft limit generally involves non-trivial

colour correlations. While it is clear that gluon splittings come with a factor of T_R, one possibility to assign gluon-emission colour factors is as follows: quark-quark emissions get a colour factor of $2C_F$, gluon-gluon antennae a factor of C_A, and quark-gluon antennae an interpolation $(C_A + 2C_F)/2$. As for dipole showers, the attribution of colour factors in this way yields incorrect subleading-N_C factors even for leading logarithms [277, 281–283, 288–291, 319–322].

It must be emphasised that antenna functions are not unique in the choice of constant and finite terms vanishing in the soft or collinear limits, *i.e.*, $y_{ij} \to 0$, $y_{jk} \to 0$. Different choices are possible, depending on whether the antenna functions are derived from physical matrix elements [245, 251, 255, 256], helicity amplitudes [323–326], or amplitude factorisation (*cf.* eqs. (3.20) and (3.21)) [193, 194, 197, 198, 217]. Regarding the latter, it should be highlighted that antenna factorisation holds on the amplitude level. Initial-final (IF) and initial-initial (II) antenna functions can be derived from eqs. (3.78) to (3.81) by crossing invariance, *i.e.*, by crossing one (IF) or two (II) final-state particles into the initial state, $p_i \to -p_i$.

From the form of eqs. (3.78) to (3.80), it is obvious that the antenna functions recover the eikonal factor eq. (3.35) in the soft limit

$$s_{ij} \to \lambda s_{ij}\,, \quad s_{jk} \to \lambda s_{jk} \quad \lambda \to 0. \tag{3.82}$$

In the $(i-j)$ collinear limit,

$$s_{ij} \to \lambda s_{ij}\,, \quad s_{jk} \to (1-z)(s_{ik}+s_{jk})\,, \lambda \to 0\,, \tag{3.83}$$

on the other hand, the antenna functions eqs. (3.78), (3.79) and (3.81) yield the full DGLAP kernels eqs. (3.59) to (3.61) only in the case of $\mathrm{q}-\mathrm{g}$ collinear limits, while in case of $\mathrm{g}-\mathrm{g}$ and $\mathrm{q}-\bar{\mathrm{q}}$ collinear limits, the collinear singularity is partitioned between two neighbouring antennae, which are related in the collinear limit by $z \leftrightarrow 1-z$. This is the case, because eqs. (3.78) to (3.81) are *global* (or *sub-*) antenna functions, constructed so that they correspond to a single antenna (colour dipole) with a unique kinematic mapping $p_I + p_K = p_i + p_j + p_k$. There exists an alternative formulation of the antenna framework, which uses so-called *sector antenna* functions [193, 194, 323–325], which reduce to both the eikonal factor eq. (3.35) in the soft limit as well as the appropriate DGLAP kernel eqs. (3.59) to (3.61) in any collinear limit. A final-state shower based on sector antennae was presented in [327]. A full-fledged implementation of sector showers in VINCIA, including initial-state radiation and coloured-resonance decays will be presented in chapter 4.

Antenna showers are ordered in a symmetric notion of transverse momentum, for final-final emissions known as the ARIADNE $p_\perp$ defined by

$$p_\perp^2 = \frac{s_{ij}s_{jk}}{s_{IK}} = y_{ij}y_{jk}s_{IK}\,, \quad s_{IK} = s_{ij}+s_{jk}+s_{ik} \tag{3.84}$$

which can be generalised to the massive case with initial-state partons as

$$p_\perp^2 = \frac{\bar{q}_{ij}\bar{q}_{jk}}{s}\,, \quad \bar{q}_{ij} = \begin{cases} s_{ij}+m_i^2+m_j^2-m_I^2 & \text{if } i \text{ final} \\ s_{ij}-m_i^2-m_j^2+m_I^2 & \text{if } i \text{ initial} \end{cases}\,, \quad s = \begin{cases} s_{ij}+s_{jk}+s_{ik} & \text{FF} \\ s_{ij}+s_{ik} & \text{IF} \\ s_{ij} & \text{II} \end{cases} \tag{3.85}$$

Following the idea of recoiler-emitter agnostic evolution, the branching recoil is shared with both parent partons at the same time, possibly constrained by one (IF) or two (II) partons remaining aligned with the beam axis. Note that this treatment is different than in dipole showers, section 3.2.2, because both initial-state partons are evolved at the same

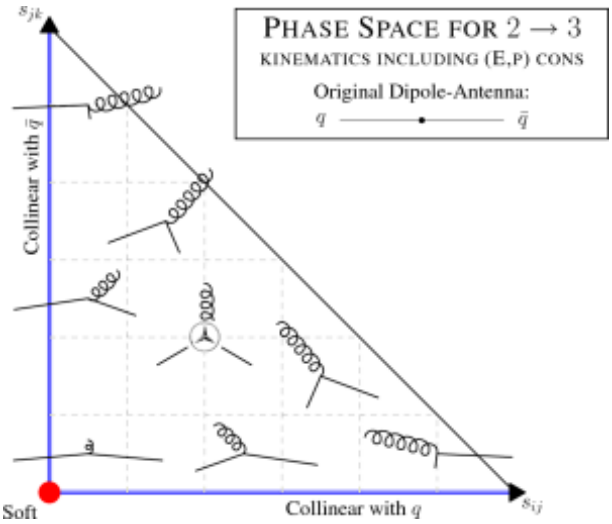


Figure 3.5: *Illustration of the recoiler-emitter agnostic antenna recoil for a final-final* $\mathrm{q}-\bar{\mathrm{q}}$ *antenna. Taken from* [277].

time in the II case, *cf.* also [242, 244]. In any case, the antenna kinematics contain the dipole recoil schemes in the limit of collinear kinematics. For FF branchings, this is visualised in fig. 3.5.

To date, the VINCIA antenna shower in PYTHIA 8.3 [15] includes initial- and final-state showers [317], showers in coloured resonance decays [328], interleaved coherent QED branchings [329, 330], and electroweak showers[16] [331].

3.3 Combining Parton Showers with Fixed-Order Calculations

Resummation (both analytical and numerical) is accurate in the regions of phase space in which emissions are soft or collinear, but becomes inaccurate in the "hard" regions of phase space, which is dominated by highly-energetic, well-separated jets. In the latter region, fixed-order calculations provide reliable and precise predictions. In other words, fixed-order calculations describe hard jets, while resummation techniques account for the jet substructure. Moreover, parton showers a priori only retain the inclusive (LO) accuracy of the Born-level calculation, while fixed-order calculations typically obtain higher (NLO, NNLO, ...) accuracy.

Given that fixed-order calculations and (parton-shower) resummation are largely complementary, it appears prudent to combine the two approaches and include higher-order corrections in parton showers. In practice, this is complicated by the fact that there exists a partial overlap, so that simply adding the two types of calculations *overcounts* emissions. This overcounting has to be avoided by a systematic combination, for which two strategies exist: *matching* and *merging*. The distinction between what is referred to as a matching scheme and what is called a merging scheme is not always clear. To disambiguate the terminology, we will use the terms in the following way:

Matching describes the combination of a fixed-order (typically NLO) calculation with a parton shower, avoiding double-counting in overlap regions (often analytically).

Merging describes the combination of multiple inclusive (N)LO event samples into a single inclusive one, dressed with additional shower radiation, accounting for Sudakov

[15] With the release of PYTHIA 8.3, VINCIA was incorporated into the main PYTHIA code and the plug-in structure discontinued.

[16] VINCIA's EW shower is based on (quasi-)collinear factorisation and not based on full antenna functions.

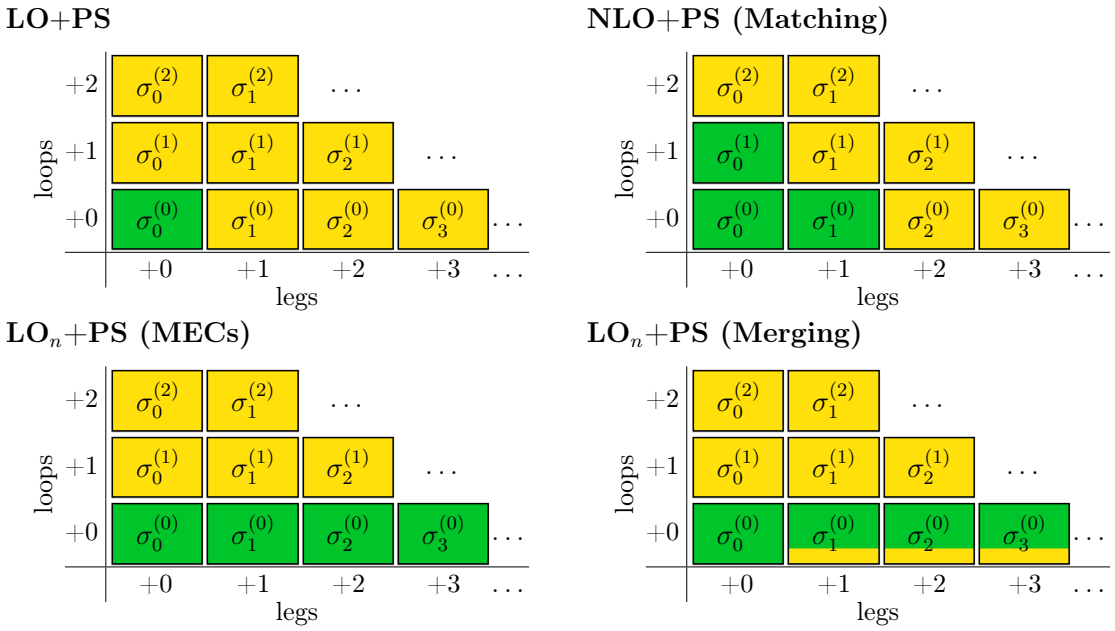


Figure 3.6: *Schematic overview of different methods to combine fixed-order calculations with parton showers. Adapted from* [327].

suppression and avoiding double-counting in overlap regions (often via phase-space slicing).

However, as will become apparent later, even this distinction is not always suitable to describe a given fixed-order-plus-parton-shower combination scheme.

An (incomplete) overview of different matching and merging techniques is schematically sketched and contrasted to the simple LO+PS treatment in fig. 3.6. The ideas behind the schemes visualised in the figure will be described in the following subsections.

3.3.1 Matrix-Element Corrections

Historically the first method was that of matrix-element corrections (MECs) [332–335], where the shower kernel itself is corrected to the full matrix element after the first emission,

$$\Pi_n^{\mathrm{MEC}}(t_n, t_{n+1}) = \exp\left\{ -\int_{t_{n+1}}^{t_n} \mathrm{d}\Phi_{+1}(t', \zeta', \phi') \, \frac{\mathrm{R}_n(\Phi_\mathrm{B}, \Phi_{+1}(t', \zeta', \phi'))}{\mathrm{B}_n(\Phi_\mathrm{B})} \right\} . \tag{3.86}$$

Including only the first emission, the expectation value of an infrared-safe observable O reads

$$\begin{aligned} \langle O \rangle_n^{\mathrm{MEC}} = & \int \mathrm{d}\Phi_\mathrm{B} \, \mathrm{B}_n(\Phi_\mathrm{B}; \mu_\mathrm{R}, \mu_\mathrm{F}) \\ & \times \left[\Pi_n^{\mathrm{MEC}}(t, t_\mathrm{c}) O(\Phi_\mathrm{B}) + \int_{t_\mathrm{c}}^{t} \mathrm{d}\Phi_{+1}(t', \zeta', \phi') \, \frac{\mathrm{R}_n(\Phi_n, \Phi_{+1})}{\mathrm{B}_n(\Phi_n)} \Pi_n^{\mathrm{MEC}}(t, t') O(\Phi_\mathrm{R}) \right] . \end{aligned} \tag{3.87}$$

It is straight-forward to see that this corrects the Born+1j prediction to the real-radiation matrix element, so that now both, the exclusive Born and Born + 1j predictions obtain LO precision. This method has later been extended to correct further emissions as well [277, 316, 317, 327, 336].

An important feature of MECs is that they rely on the parton shower to generate the (Sudakov-weighted) phase space of all multiplicities, so avoid the need of non-trivial and inefficient high-multiplicity phase space generators, *cf.* section 3.1.1. This does not come without a price, however, since conventional showers are at most[17] able to fill the phase space of the first emission due to the strong ordering criterion, *cf.* section 3.2. Matrix-element-corrected showers therefore systematically miss potentially important higher-multiplicity corrections. Related to this, there is a further subtlety in the case of initial-state radiation, where the factorisation theorem, *cf.* section 2.3, implies that the *resummation scale*, *i.e.*, the scale at which showers are started, should be the factorisation scale μ_F. Again due to the strong ordering condition, a "wimpy" shower starting at the factorisation scale is incapable of filling the full phase space, even of the first emission. This leads to visible effects *e.g.* in the transverse-momentum distribution of the Z^0 boson, where the high-p_T tail is entirely missed, *cf. e.g.* [308]. So-called "power showers" [337] avoid this by letting the shower start at the phase space maximum. This, however, extends the resummation region beyond its applicability in disagreement with the factorisation theorem and introduces spurious PDF ratios of the form $f_{a/h_1}(x_a,\mu_\mathrm{PS})/f_{a/h_1}(x_a,\mu_\mathrm{F})$.

3.3.2 NLO Matching

Two general and widely applied NLO matching schemes have been developed, namely MC@NLO [338] and POWHEG [339–341], with the former being automated in the SHERPA [342] and MADGRAPH_AMC@NLO [115] event-generation frameworks and the latter available through the POWHEGBOX program [229] [18].

The first method to match showers to NLO calculations has been described for the case of final-state radiation in [345] and is since routinely employed for resonance decays in PYTHIA. Today, this scheme is known under the name of POWHEG and extends MECs, *cf.* eq. (3.86), to NLO accuracy in the Born,

$$\begin{aligned}\langle O\rangle_n^{\text{POWHEG}} = \int \mathrm{d}\Phi_\mathrm{B}\, k_\mathrm{NLO}^{\text{POWHEG}}(\Phi_\mathrm{B};\mu_\mathrm{R},\mu_\mathrm{F})\mathrm{B}_n(\Phi_\mathrm{B};\mu_\mathrm{R},\mu_\mathrm{F})\Bigg[\Pi_n^{\text{MEC}}(t,t_\mathrm{c})O(\Phi_\mathrm{B})\\ +\int_{t_\mathrm{c}}^{t}\mathrm{d}\Phi_{+1}(t',\zeta',\phi')\frac{\mathrm{R}_n(\Phi_\mathrm{B},\Phi_{+1})}{\mathrm{B}_n(\Phi_\mathrm{B})}\Pi_n^{\text{MEC}}(t,t')O(\Phi_\mathrm{R})\Bigg]\end{aligned} \tag{3.88}$$

by introducing a Born-differential NLO K-factor

$$\begin{aligned}k_\mathrm{NLO}^{\text{POWHEG}}(\Phi_\mathrm{B};\mu_\mathrm{R},\mu_\mathrm{F}) = 1+\frac{\mathrm{V}_n(\Phi_\mathrm{B};\mu_\mathrm{R},\mu_\mathrm{F})}{\mathrm{B}_n(\Phi_\mathrm{B};\mu_\mathrm{R},\mu_\mathrm{F})}+\frac{\mathrm{I}_n(\Phi_\mathrm{B};\mu_\mathrm{R},\mu_\mathrm{F})}{\mathrm{B}_n(\Phi_\mathrm{B};\mu_\mathrm{R},\mu_\mathrm{F})}\\ +\int\mathrm{d}\Phi_{+1}\left[\frac{\mathrm{R}_n(\Phi_\mathrm{B},\Phi_{+1};\mu_\mathrm{R},\mu_\mathrm{F})}{\mathrm{B}_n(\Phi_\mathrm{B};\mu_\mathrm{R},\mu_\mathrm{F})}-\frac{\mathrm{S}_n(\Phi_\mathrm{B},\Phi_{+1};\mu_\mathrm{R},\mu_\mathrm{F})}{\mathrm{B}_n(\Phi_\mathrm{B};\mu_\mathrm{R},\mu_\mathrm{F})}\right].\end{aligned} \tag{3.89}$$

While the POWHEG method is guaranteed to yield positive-definite event weights as long as the NLO correction is smaller than the LO prediction, there are two noteworthy subtleties. Just as for tree-level MECs, eq. (3.86), the POWHEG method requires the parton shower to generate the one-emission phase space, extending the parton-shower resummation to the region $t > \mu_\mathrm{F}^2$. Furthermore, using the ratio R/B as shower branching kernels exponentiates process-dependent finite terms in the no-branching probability, which may

[17] Angular-ordered parton showers contain a dead zone even in the phase space of the first emission.

[18] The POWHEG method was implemented in previous versions of the SHERPA framework as well [341], but has been discontinued starting from the 2.0 series [342]. Similar multiplicative matching strategies are also implemented in WHIZARD [343] and HERWIG/MATCHBOX [120, 344].

be regarded as inelegant since it makes the shower resummation process-dependent. In chapter 7, a method extending the POWHEG scheme to NNLO is presented.

The above-mentioned subtleties are not present in MC@NLO, because the real correction is divided into an infrared-singular (soft) and an infrared-finite (hard) part[19], $\mathrm{R} = \mathrm{R}^{(\mathrm{S})} + \mathrm{R}^{(\mathrm{H})}$ [342],

$$\begin{aligned}\langle O\rangle_n^{\text{MC@NLO}} = \int \mathrm{d}\Phi_\mathrm{B}\, \bar{\mathrm{B}}_n(\Phi_\mathrm{B};\mu_\mathrm{R},\mu_\mathrm{F}) \Bigg[&\Pi_n^{(\mathrm{S})}(t,t_\mathrm{c})O(\Phi_\mathrm{B}) \\ &+ \int_{t_\mathrm{c}}^{t} \mathrm{d}\Phi_{+1}(t',\zeta',\phi')\frac{\mathrm{R}_n^{(\mathrm{S})}(\Phi_n,\Phi_{+1})}{\mathrm{B}_n(\Phi_n)}\Pi_n^{(\mathrm{S})}(t,t')O(\Phi_\mathrm{R})\Bigg] \\ &+ \int \mathrm{d}\Phi_\mathrm{R}\frac{\mathrm{R}_n^{(\mathrm{H})}(\Phi_n,\Phi_{+1})}{\mathrm{B}_n(\Phi_n)}O(\Phi_\mathrm{R})\end{aligned} \tag{3.90}$$

with

$$\begin{aligned}\bar{\mathrm{B}}_n(\Phi_\mathrm{B};\mu_\mathrm{R},\mu_\mathrm{F}) = \mathrm{B}_n(\Phi_\mathrm{B};\mu_\mathrm{R},\mu_\mathrm{F}) + \mathrm{V}_n(\Phi_\mathrm{B};\mu_\mathrm{R},\mu_\mathrm{F}) + \mathrm{I}_n(\Phi_\mathrm{B};\mu_\mathrm{R},\mu_\mathrm{F}) \\ + \int \mathrm{d}\Phi_{+1}\,\left[\mathrm{R}_n^{(\mathrm{S})}(\Phi_\mathrm{B},\Phi_{+1};\mu_\mathrm{R},\mu_\mathrm{F}) - \mathrm{S}_n(\Phi_\mathrm{B},\Phi_{+1};\mu_\mathrm{R},\mu_\mathrm{F})\right]\end{aligned} \tag{3.91}$$

and

$$\Pi_n^{(\mathrm{S})}(t_n,t_{n+1}) = \exp\left\{-\int_{t_{n+1}}^{t_n}\mathrm{d}\Phi_{+1}(t',\zeta',\phi')\,\frac{\mathrm{R}_n^{(\mathrm{S})}(\Phi_\mathrm{B},\Phi_{+1}(t',\zeta',\phi'))}{\mathrm{B}_n(\Phi_\mathrm{B})}\right\} \tag{3.92}$$

It must be noted that due to the intimate interplay between the shower branching kernels and the NLO subtraction terms, the MC@NLO method is shower-dependent and has to be implemented explicitly for every shower algorithm it is supposed to be used with. Related to this, it should also be mentioned that practical implementations differ in precisely the treatment of the $\mathrm{R}_n^{(\mathrm{S})} - \mathrm{S}_n$ term. While the SHERPA (S-MC@NLO) implementation amends the shower by subleading-colour branchings (with negative weights) to match the infrared singularity structure of $\mathrm{R}_n^{(\mathrm{S})}$ exactly, this is avoided in the MADGRAPH_AMC@NLO (AMC@NLO) implementation.

3.3.3 Merging

In multi-jet merging, multiple event samples are combined into a single inclusive sample. In order for this to be done safely without double counting any emissions, the initially inclusive event samples first have to be made exclusive and shower emissions have to be constrained to the resummation region. In the CKKW [346,347] and derived schemes, the latter is achieved by the introduction of a (somewhat arbitrary) *merging scale* k_MS, which separates the shower region ($k < k_\mathrm{MS}$) from the matrix-element region ($k > k_\mathrm{MS}$). Up to the first emission, merging an n-jet and $(n+1)$-jet calculation thus takes the schematic form

$$\langle O\rangle_n^{\text{ME}_1\text{PS}} = \int \mathrm{d}\Phi_{\mathrm{B}_n}\,\mathrm{B}_n(\Phi_\mathrm{B};\mu_\mathrm{R},\mu_\mathrm{F})\Bigg[\Pi_n(t_n,t_\mathrm{c})O(\Phi_{\mathrm{B}_n}) \tag{3.93}$$

[19] A similar (but not identical) division of the real correction into a hard and soft part is possible also in the POWHEG scheme via the "`hdamp`" parameter.

$$+\int \mathrm{d}\Phi_{+1}\,\mathrm{K}_{n\mapsto n+1}(\Phi_{\mathrm{B}_n},\Phi_{+1})\Pi_n(t_n,t_{n+1})\theta(k_{\mathrm{MS}}-k_{n+1})O(\Phi_{\mathrm{B}_{n+1}})\Big]$$
$$+\int \mathrm{d}\Phi_{\mathrm{B}_{n+1}}\,\mathrm{B}_{n+1}(\Phi_{\mathrm{B}_{n+1}};\mu_{\mathrm{R}},\mu_{\mathrm{F}})\tilde{\Pi}_n(t_n,t_{n+1})\theta(k_{n+1}-k_{\mathrm{MS}})O(\Phi_{\mathrm{B}_{n+1}})\,,$$

with the Heaviside step function θ. The first two lines describe the Born state including the effect of a *truncated* and *vetoed* shower, *i.e.*, a shower which is truncated after the first emission with branchings above the merging scale vetoed. The interesting part is the third line, which describes the addition of the hard matrix element in the region above the merging scale cut k_{MS}. While it is clear that states B_{n+1} with $k_{n+1} > k_{\mathrm{MS}}$ can be excluded from the calculation by simply calculating the scale k_{n+1} and vetoing the state if the result falls below the merging scale cut, the meaning of the no-branching probability $\tilde{\Pi}_n$ is a priori undefined, because this state was not generated by the parton shower. To make this meaningful, a *shower history* is defined, *i.e.*, a sequence of on-shell configurations,

$$\{S_{\mathrm{Born}}, S_{\mathrm{Born+1}}, S_{\mathrm{Born+2}}, \ldots, S_{\mathrm{Born}+m}\}$$

which corresponds to a sequence of the parton-shower evolution variable $\{t_0, t_1, t_2, \ldots, t_m\}$. This procedure now allows to define no-branching probabilities $\Pi_m(\Phi_{\mathrm{B}_m}; t_m, t_{m+1})$ between every adjacent pair of nodes in the shower history. It is the construction of the shower history and the treatment of the no-branching probability where different merging approaches differ. In the original CKKW scheme, the history was simply constructed using a k_{T} clustering algorithm and analytic NLL Sudakov factors used as $\tilde{\Pi}_n$. This is not ideal because of three reasons: in general, the shower evolution variable will not be identical to the k_{T} variable used in the clustering algorithm; similarly, the no-branching probabilities in the shower are not identical to the analytically calculated ones; lastly, the shower will also generate sequences which do not correspond to the ones defined in the deterministic clustering algorithm. Taken together, these three issues can potentially lead to overcounted or missed regions in phase space. The first two problems have been overcome by various refinements of the CKKW algorithm. To this end, the CKKW-L method [348–350] and MeTs [351] account for the mismatch between the shower ordering variable and the variable in terms of which the merging scale is defined and use numerical no-branching probabilities, generated by truncated showers. The latter point ensures that $\tilde{\Pi}_m(\Phi_{\mathrm{B}_m}; t_m, t_{m+1}) = \Pi_m(\Phi_{\mathrm{B}_m}; t_m, t_{m+1})$ in eq. (3.93), *i.e.*, that the same no-branching probabilities are used in the shower and the merging. In this context, truncated showers are showers that are started on any node S_m and are stopped after the first emission has been generated. A subtlety is, however, still present in the definition of the shower history. In the CKKW-L algorithm, a proper shower history is generated by considering all possible branching sequences the shower at hand may have generated to arrive at the given hard configuration S_n. Then, the most likely sequence is picked by maximising over the product of shower branching kernels. While this is guaranteed to generate Sudakov factors in exactly the same way the shower would have done, this procedure quickly becomes cumbersome, as the number of shower paths grows at least factorially with the number of legs. A method that overcomes this problem of CKKW-L is presented in chapter 5.

In the CKKW-L algorithm, multiple event samples are combined according to the following formula

$$\langle O\rangle_n^{\text{CKKW-L}} = \int \mathrm{d}\Phi_{\mathrm{B}_n}\, w_n \mathrm{B}_n(\Phi_{\mathrm{B}_n};\mu_{\mathrm{R}},\mu_{\mathrm{F}})O(\Phi_{\mathrm{B}_n})$$
$$+\int \mathrm{d}\Phi_{\mathrm{B}_{n+1}}\, w_{n+1}\mathrm{B}_{n+1}(\Phi_{\mathrm{B}_{n+1}};\mu_{\mathrm{R}},\mu_{\mathrm{F}})\theta(k_{n+1}-k_{\mathrm{MS}})O(\Phi_{\mathrm{B}_{n+1}})$$

$$+\dots$$
$$+\int \mathrm{d}\Phi_{\mathrm{B}_{n+k}}\, w_{n+k}\mathrm{B}_{n+k}(\Phi_{\mathrm{B}_{n+k}};\mu_\mathrm{R},\mu_\mathrm{F})\theta(k_{n+k}-k_\mathrm{MS})O(\Phi_{\mathrm{B}_{n+k}})\,, \tag{3.94}$$

where the weights w_m reweight inclusive events to exclusive ones and ensure a smooth merging with the shower by additional α_S and PDF weights,

$$w_m = \frac{f_{a_m/h_1}(x_m,t_m)}{f_{a_m/h_1}(x_m,\mu_\mathrm{F}^2)}\frac{f_{b_m/h_2}(x_m,t_m)}{f_{b_m/h_2}(x_m,\mu_\mathrm{F}^2)} \times\prod_{i=0}^{m-1}\frac{\alpha_{\mathrm{s,PS}}(t_{i+1})}{\alpha_{\mathrm{s,ME}}}\frac{f_{a_i/h_1}(x_i,t_i)}{f_{a_i/h_1}(x_i,t_{i+1})}\frac{f_{b_m/h_2}(x_i,t_i)}{f_{b_m/h_2}(x_i,t_{i+1})}\Pi_i(t_i,t_{i+1})\,. \tag{3.95}$$

It is straight-forward to see from eq. (3.94) that the CKKW-L scheme changes the inclusive Born-level cross section by the addition of (weighted) higher-multiplicity samples. The inclusive cross section of the Born-level event sample can be restored by so-called unitarised merging schemes (UMEPS) [352, 353] via the explicit subtraction of higher-multiplicity events from lower-multiplicity events just as the shower operator eq. (3.56) dictates it.

As NLO calculations have become widely available and experimental precision increases, multiple refinements of the above LO merging schemes have been developed to extend it to NLO. The first NLO merging scheme, called NL3, was developed in [354] and is based on CKKW-L. The MENLOPS scheme [355, 356] extends the METS scheme to combine an NLO calculation for the Born multiplicity with multiple LO calculations at higher multiplicities and has been further refined to include NLO calculations at higher multiplicities under the name MEPS@NLO [357, 358]. The UNLOPS scheme [359, 360] may either be viewed as an NLO extension of UMEPS or a unitarised version of NL3. Within UNLOPS, NLO accuracy is obtained for exclusive higher-multiplicity predictions, while the inclusive NLO cross section of the lowest-multiplicity sample is retained.

It may be viewed as inelegant that merging schemes require the introduction of an additional parameter, the merging scale. In the method presented in [336], dubbed MOPS, this is avoided by a judicious combination of the CKKW-L algorithm with a matrix-element corrected parton shower, *cf.* section 3.3.1. In that scheme, higher-order matrix elements are only merged into the shower when the associated particle configuration cannot be produced by a strongly-ordered shower (so-called *incomplete histories*). As such, no merging scale is required to decide when the fixed-order calculation is ought to be used over the shower approximation. An alternative CKKW-based scheme has been developed under the name MINLO [361, 362] as an extension of the POWHEG matching scheme, *cf.* section 3.3.2. It avoids the introduction of a merging scale in the NLO merging via the use of analytic Sudakovs, which regularise singularities in higher-multiplicities, so that these can be performed at very low resolution scales.

A tree-level merging algorithm that does not fall into the category of the ones described so far is the MLM scheme [363, 364], which utilises a simple jet-matching[20] algorithm to remove the radiation overlap, *i.e.*, checks whether new hard jets compared to the Born process have been generated after the shower evolution has stoppped. It has been extended to NLO in the FxFx scheme [365], using analytic Sudakov weights akin to the CKKW and MINLO methods.

Because of the way merging schemes avoid singularities pertaining to higher-order matrix elements (either through a merging scale or Sudakov weights), it has been possible

[20] The term "matching" here is not to be confused with the term used in NLO matching strategies.

to define merging-based NNLO matching schemes. Instead of a differential matching of the shower kernels with the fixed-order calculation, such approaches merge NNLO-accurate Born calculations with NLO Born + 1-jet calculations and LO Born + $2j$ calculations, utilising available (NLO) merging techniques. The UN2LOPS scheme [366–369] extends the UNLOPS merging to the second order, while the MINNLO$_{\text{PS}}$ [370, 371] and other approaches [361, 362, 372–377] extend the MINLO procedure. Recently, a proof of concept for an UNLOPS-based N^{3}LO+PS matching strategy has been presented in [378].

3.4 Hadronisation

Once the particles in the simulation reach a scale of $\Lambda_{\text{QCD}} \approx 1$ GeV, perturbative QCD is no longer applicable due to the growing value of the strong coupling at low energies, *cf.* fig. 2.1, and the parton-shower evolution has to be stopped. The simulation then transfers to the *hadronisation* or *fragmentation* phase, during which the transition from coloured *partons* to composite, colour-singlet *hadrons* is made. If unstable hadrons are formed, they decay further into stable particles afterwards. As a perturbative treatment of this process is prohibited, these hadronisation models cannot be derived from first principles but have to be based on judicious, QCD-inspired assumptions. Hadronisation models are therefore intrinsically *non-perturbative*.

Among the first hadronisation models was the Field-Feynman model [379, 380] based on *independent fragmentation* [381], meaning that partons fragment into hadrons independently of each other. It modelled fragmentation as iterative branchings $\mathrm{q}_1 \mapsto \mathrm{h}_1\mathrm{q}_2$ with a Gaussian transverse momentum distribution for each branching, disrespecting four-momentum conservation. Despite its simplistic nature, it facilitated the construction of early Monte Carlo generators for $\mathrm{e}^+\mathrm{e}^-$ collisions [382, 383]. To date, two paradigms build the foundations of fragmentation modules in Monte Carlo simulations: the (Lund) string model implemented in PYTHIA [384–388] and the cluster fragmentation model implemented in HERWIG [389] and SHERPA [390]. These two models are illustrated in fig. 3.7 and will be discussed briefly in the following two subsections.

3.4.1 String Fragmentation

As the basis of the early Lund Monte Carlos [91–97], the string model [384–388] is still one of the cornerstones of modern-day PYTHIA [101, 102, 391]. While the Lund string model is the best known, other string-based fragmentation models have been developed [392–396]. In fact, the Artru-Mennessier model [394, 395], which already provided a Monte Carlo implementation, was the first-ever hadronisation model (even before the Field-Feynman one mentioned above), but went mostly unnoticed.

String fragmentation is based on the *linear confinement* property of QCD, which describes the potential between a quark and an antiquark at large spatial separation by a *flux tube*, meaning a confined "tube" of colour-field lines, as indicated by lattice QCD studies [397–400]. The potential energy between two static quarks is thus described by a Cornell potential [401],

$$V(r) = \kappa r - \frac{a}{r}\,, \tag{3.96}$$

which behaves like a Coulomb potential at small separation $r \to 0$, while approximating a linear potential with increasing separation $r \gtrsim 1$ fm. The parameter κ is known as the *string tension* with a value of about 1 GeV/fm. In the practical implementation of the Lund string model in PYTHIA, the Coulomb term $\frac{a}{r}$ is ignored, as it is believed to be less important for the hadron creation process but rather for the internal structure of hadrons.

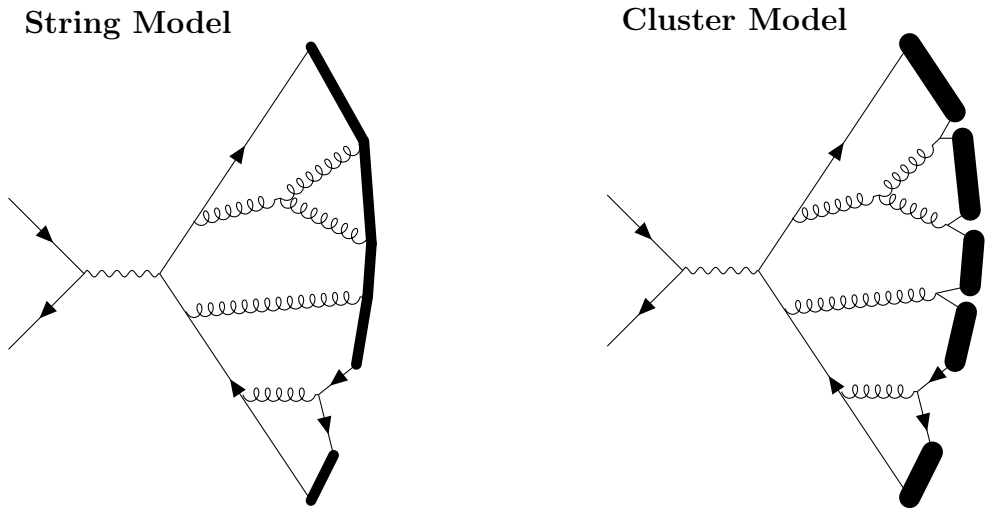


Figure 3.7: *Illustration of the string (left) and cluster (right) hadronisation models in* $e^+e^- \to$ jets. *Note that in the latter gluons are split into quarks just before the cluster formation. Inspired by* [116].

While being derived from considering the QCD potential between a quark-antiquark pair, it is simple to incorporate gluons in the string model. In momentum space, these are represented by "kinks" in the string. With this picture in mind, it is simple to see that string fragmentation is infrared safe: for a $q - g - \bar{q}$ string, the kink vanishes if the gluon has infinitely soft momentum or is infinitesimally close to one of the string ends, reducing the string to a $q - \bar{q}$ one[21].

In the string model, hadrons are created by *string breaks* on average along a hyperbola in space-time. These occur as the two quarks move away from each other, such that the energy stored in the string becomes large enough to create a new quark-antiquark pair. These are created in a single point and the quarks then quantum-mechanically tunnel to the classically-allowed region, allowing them to obtain transverse momentum, which is created from the field energy between them. Following the rationale of e^+e^- pair creation in a constant electric field within the WKB approximation [402], this process happens with probability

$$P \propto \exp\left(-\frac{\pi m_\perp^2}{\kappa}\right) = \exp\left(-\frac{\pi p_\perp^2}{\kappa}\right) \exp\left(-\frac{\pi m_q^2}{\kappa}\right) , \tag{3.97}$$

which implies that heavier quarks are less likely to be produced by string breaks. From the space-like separation and the constant string tension it follows that every two string breaks are spacelike separated, *i.e.*, causally disconnected. The string breaks are, however, connected by the constraint that the invariant mass of the string between the two breaks corresponds to an on-shell hadron. There is no particular order in which hadrons should be created from a string, so that string breaks can be simulated iteratively. Requiring that an iteration starting on the q side yields the same answer as one starting on the $\bar{q}$ side, gives the *Lund symmetric fragmentation function* [385],

$$f(z) \propto \frac{(1-z)^a}{z} \exp\left(-\frac{bm_\perp^2}{z}\right) , \tag{3.98}$$

with free parameters a and b. It describes the distribution of the light cone momentum fraction z taken by the new hadron from the string (which is left with a fraction $1 - z$).

[21] While this is true physically, there might still be numerical issues related to this in practice if the string contains too many adjacent soft gluons.

Baryon formation proceeds in three ways in the string model. Firstly, instead of $\mathrm{q}-\bar{\mathrm{q}}$ pairs, occasionally so-called diquarks, *i.e.*, $\mathrm{q}-\mathrm{q}$ states, in an (anti-)triplet representation are pair-produced in string breaks [403, 404],

$$(\mathrm{q}_1\bar{\mathrm{q}}_1) \mapsto (\mathrm{q}_1(\mathrm{q}_2\mathrm{q}_3)(\bar{\mathrm{q}}_3\bar{\mathrm{q}}_2)\bar{\mathrm{q}}_1) = (\mathrm{q}_1\mathrm{q}_2\mathrm{q}_3)(\bar{\mathrm{q}}_3\bar{\mathrm{q}}_2\bar{\mathrm{q}}_1)\,.$$

Secondly, baryons can be produced by what is known as the "popcorn" mechanism [405, 406], in which no diquarks are needed. Instead, a new $\mathrm{q}\bar{\mathrm{q}}$ is produced in such a way that it does not break the string, which only happens upon a secondary quark-pair creation, *i.e.*,

$$(\mathrm{q}_1\bar{\mathrm{q}}_1) \mapsto (\mathrm{q}_1(\mathrm{q}_2\bar{\mathrm{q}}_2)\bar{\mathrm{q}}_1) \mapsto (\mathrm{q}_1(\mathrm{q}_2(\mathrm{q}_3\bar{\mathrm{q}}_3)\bar{\mathrm{q}}_2)\bar{\mathrm{q}}_1) = (\mathrm{q}_1\mathrm{q}_2\mathrm{q}_3)(\bar{\mathrm{q}}_3\bar{\mathrm{q}}_2\bar{\mathrm{q}}_1)\,.$$

Thirdly, vertices in which three strings meet, so-called string junctions, can lead to the production of baryons. String junctions exist due to the presence of a non-vanishing Levi-Civita symbol in QCD.

In the context of high-energy hadron collisions, we will rarely (if at all) encounter events in which only a single colour string is present. While the model was originally developed with such a situation in mind (which might be a good approximation *e.g.* for electron-positron collisions), collective effects are taken into account by various refinements of the original model, *cf. e.g.* [407–413].

3.4.2 Cluster Fragmentation

To date, models based on cluster fragmentation provide the default hadronisation modules of the HERWIG and SHERPA event generators [389, 390], with slightly different specific implementations. Historically the first cluster hadronisation models were developed and implemented in Monte Carlo codes by Fox, Field, and Wolfram [414,415] following the idea of the local parton-hadron duality [416, 417], implying that the distribution of final-state hadrons follows closely the one of partons at the end of a parton cascade.

Cluster fragmentation is based on a property of QCD known as *preconfinement* [418], which states that the mass distribution of colour-singlet "clusters" at the end of a parton cascade decouples from all other scales in the event (except Λ_{QCD}) once a (small) "preconfinement scale" is reached. In particular, the cluster mass distribution is independent of the resummation scale at which the parton shower has been started and has a sharp peak at low masses. Because of this, cluster hadronisation is initiated by non-perturbative gluon splittings, $\mathrm{g} \mapsto \mathrm{q}\bar{\mathrm{q}}$, *cf.* fig. 3.7. They ensure that only colour-anticolour pairs are present in the event, from which then colour-singlet clusters are produced by combining adjacent colour pairs. This is a sensible procedure, as in the large-N_{C} limit, adjacency in colour implies adjacency in phase space, especially for angular-ordered showers where branchings become decreasingly close in angle at the end of the shower. Because of the way they are formed, the initial clusters will have a phase-space distribution similar to the partons after the shower evolution has ceased. As such, the imposed gluon splittings may be seen as a direct implementation of the local parton-hadron duality mentioned above.

As a consequence of the procedure to enforce gluon splittings to start the fragmentation process, gluons have to be assigned a *constituent mass* of around $m_{\mathrm{g}} \approx 1$ GeV. Because strange quarks have a larger constituent mass, $m_{\mathrm{s}} \approx 450$ MeV, than up and down quarks, which are chosen roughly as a third of the proton mass $m_{\mathrm{u/d}} \approx 330$ MeV, the flavour distribution in clusters is biased towards the lighter quarks, with non-perturbative gluon splittings to heavy flavours (charms and bottoms) entirely forbidden.

Baryons can be produced either via cluster decays to baryon-antibaryon pairs or via

non-perturbative gluon decays to diquarks, $g \mapsto (q_1 q_2)(\bar{q}_1 \bar{q}_2)$ [419]. In addition to cluster decays to hadrons, heavy clusters may undergo *cluster fissioning* beforehand, allowing them to produce lighter clusters by quark-pair creation, similar in spirit to string breaks in the string model. During this phase, strangeness production may be enhanced [420] to increase agreement with experimental observations as *e.g.* by the ALICE collaboration, *cf.* [421].

4

Sector Showers for Hadron Collisions

Parton shower algorithms are iterative, *i.e.*, starting from the Born configuration, further emissions are generated successively according to branching probabilities derived from the soft and collinear limits, *cf.* eqs. (3.33) and (3.34). As multiple partons exist on the Born level (at least two if any QCD evolution is to be performed) and each may emit bremsstrahlung, this evolution is not unique but probabilistic. In an n-parton configuration, all possible branchings transforming this state into a state with an additional parton, $n \mapsto n+1$, compete for the next emission scale t_{n+1} and the one with the highest scale defines the evolution step that is performed. This procedure generates a self-similar branching tree, which is inherent to all common parton shower models, *cf.* section 3.2.

Looking at such a fractal tree with n branches, it is generally not possible anymore to tell from which $(n-1)$-parton configuration it was generated. The inversion of the evolution up to a configuration with n partons thus gives rise to $\sim n!$ paths of the form

$$n \mapsto (n-1) \mapsto (n-2) \mapsto \cdots \mapsto 2. \tag{4.1}$$

This means that conventional shower algorithms have a factorially growing number of so-called shower histories, *i.e.*, sequences of intermediate states leading to a given configuration. As far as the shower evolution itself is concerned, this does not pose any problems, but shower histories are an important input for matching and merging schemes, *cf.* sections 3.3.1 to 3.3.3. Accounting for a large number of histories may therefore heavily impede efficient higher-accuracy calculations for processes with many particles.

The situation is different in so-called deterministic jet algorithms. Given an n-particle configuration, the purpose of a jet algorithm is to define well-separated objects called jets, corresponding to a collimated bundle of contiguous particles. A typical way to achieve this is to define a measure of proximity and a resolution scale in terms of this measure, so that a pair of particles is clustered to a single particle whenever their measure of proximity falls below the resolution scale. If always those two particles in a configuration are clustered first that have the lowest resolution, this defines a deterministic algorithm to generate the sequence in eq. (4.1). In contrast to parton-shower algorithms, which are often based on on-shell $2 \mapsto 3$ mappings, jet-clustering algorithms utilise off-shell $2 \mapsto 1$ recombinations. Acting on a shower-generated n-parton configuration with a deterministic jet algorithm will therefore generally not yield the correct shower history.

An alternative to the conventional dipole/antenna framework, *cf.* sections 3.1.4, 3.2.2 and 3.2.3, is the sector-antenna formalism [193,194,323–325]. Within this framework, the

branching phase space is decomposed into non-overlapping sectors, each of which corresponds to a certain, well-defined branching. This phase-space decomposition allows for the definition of a shower-based jet algorithm, which exactly inverts the shower kinematics in an on-shell $3 \mapsto 2$ clustering. Such a jet algorithm was proposed long time ago in the context of the ARIADNE Lund-dipole/antenna shower [422].

In the publication included in section 4.1, the implementation of a full-fledged sector shower in the VINCIA antenna shower in PYTHIA 8.3 will be presented. The implementation builds upon an early proof-of-concept final-state sector shower in VINCIA [327] and promotes VINCIA's initial-state showers [317] and coloured-resonance decay shower [328] to sector showers.

4.1 Published Material

Sector Showers for Hadron Collisions

Helen Brooks, **Christian T Preuss**, and Peter Skands

Published in The Journal of High Energy Physics 07 (2020) 032
DOI: 10.1007/JHEP07(2020)032
e-Print: arXiv:2003.00702 [hep-ph]

Publication begins overleaf.

Published for SISSA by Springer

Received: *March 11, 2020*
Revised: *May 18, 2020*
Accepted: *June 8, 2020*
Published: *July 6, 2020*

Sector showers for hadron collisions

Helen Brooks, Christian T Preuss and Peter Skands

School of Physics and Astronomy, Monash University, Wellington Road, Clayton, VIC-3800, Australia

E-mail: christian.preuss@monash.edu, peter.skands@monash.edu

Abstract: In conventional parton showers (including ones based on dipoles/antennae), a given (Born+m)-parton configuration can typically be reached via $\mathcal{O}(m!)$ different "shower histories". In the context of matrix-element-correction and merging procedures, accounting for these histories mandates fairly complex and resource-intensive algorithms. A so far little-explored alternative in the shower context is to divide the branching phase spaces into distinct "sectors", each of which only receives contributions from a single branching kernel. This has a number of consequences including making the shower operator bijective; i.e., each parton configuration now has a *single* unique "inverse". As a first step towards developing a full-fledged matrix-element-correction and merging procedure based on such showers, we here extend the sector approach for antenna showers to hadron-hadron collisions, including mass and helicity dependence.

Keywords: QCD Phenomenology, Phenomenological Models

ArXiv ePrint: 2003.00702

https://doi.org/10.1007/JHEP07(2020)032

JHEP07(2020)032

Contents

JHEP07(2020)032

1 Introduction

General-purpose Monte Carlo event generators such as HERWIG [1], SHERPA [2], and PYTHIA [3] have become indispensable tools to study the complex event structures produced in highly energetic hadronic collisions. At the heart of these tools, parton showers resum all leading-logarithmic (LL) terms in the perturbative expansion of QCD, thereby connecting the perturbative, high-energy hard scale with the intrinsically non-perturbative,

low-energy soft scale, at which the strong coupling becomes large enough to invalidate perturbation theory.

Early parton showers were built upon the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) splitting kernels [4–6], which are suitable for describing collinear (small-angle) radiation off a hard partonic leg, but which fail to describe the interference of soft wide-angle radiation. To incorporate coherent (wide-angle) emissions in the DGLAP formalism, one can either order emissions decreasing in angle, cf. [7], or veto emissions which are not ordered in angle, cf. [8].

Alternatively, coherence effects can be directly taken into account (to leading power in $1/N_C^2$) by considering colour dipoles, that is by considering particle emissions as stemming from colour-anticolour pairs. The first shower model based on this concept was ARIADNE [9, 10], with VINCIA following a similar approach many years later [11, 12]. In the meantime, an alternative type of dipole showers, based on the Catani-Seymour (CS) approach [13, 14], also appeared [15–19]. We will not give a complete comparison between all of these approaches here. It will suffice to say that:

- For final-state showers, the original formulation of the Lund dipole model [9] (as implemented in ARIADNE) is equivalent to the "global" antenna subtraction and antenna shower framework (as implemented in VINCIA) [20].
- For initial-state showers, VINCIA is based on a backwards-evolution paradigm [21, 22], whereas ARIADNE uses a forwards-evolution picture.
- In the Lund-dipole/antenna formalisms, the soft singularity associated with each parton pair is captured by a single "antenna function" and an associated kinematics map in which both of the parents in general share the (transverse and longitudinal) recoil. In the CS dipole formalism, the radiation from each colour-connected parton pair is instead partitioned into two separate terms, in which each of the parents in turn act as the "emitter" while the other, the "recoiler", recoils purely longitudinally along the dipole axis.
- In all of the dipole/antenna formalisms, the collinear $g \mapsto gg$ singularity is partitioned into two terms, one for each colour-connected partner of the parent gluon. This partitioning is done such that the $g \mapsto gg$ DGLAP kernel is obtained by summing over the two neighbouring dipoles/antennae. The exact form of the partitioning factor differs between different models/implementations.

To improve the precision away from the collinear and soft limits, showers are usually matched or merged to fixed-order calculations. Examples of well-established approaches include MC@NLO [23] and POWHEG [24, 25] in the case of the former, and CKKW(-L) [26–28], MLM [29, 30], in the case of the latter. In recent years there has been a proliferation of refinements to merging algorithms, such as UMEPS [31], UNLOPS [32], MEPS@NLO [33, 34], FxFx [35], MINLO [36], and MINNLOPS [37].

Matching algorithms correct the hardest emission of the parton shower to the stated fixed-order accuracy, typically next-to-leading order (NLO). Merging algorithms instead

JHEP07(2020)032

Number of histories for n branchings							
	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$
CS Dipole	2	8	48	384	3840	46080	645120
Global Antenna	1	2	6	24	120	720	5040

Table 1. The number of possible shower/clustering histories (ordered + unordered) that can contribute to a given (colour-ordered) parton configuration, after n branchings starting from a single colour-anticolour pair.

combine inclusive event-samples, each of a given accuracy but of increasing multiplicity. Merging is typically favoured where it is desirable to describe multijet final states. To consistently combine multiple event samples with the parton shower, it is necessary to reweight each event with a Sudakov factor, which in turn requires obtaining a sequence of scales that represents the "parton shower history". It is possible to obtain these scales through "winner-takes-it-all" clustering methods, as is done in the case of CKKW. However, this does not correspond to a direct inversion of the parton shower and potentially results in missed areas of phase space. Ideally one would obtain all paths of possible clusterings back to the Born topology (and selecting one with its relative probability) as in CKKW-L. This, however, comes at the price of a significant computational overhead, rendering such calculations intractable for many legs (see, e.g., [38–40]).

Owing to the very formulation of both DGLAP and dipole showers, namely that every colour charge in the event can radiate another parton, successive radiation leads to a proliferation of terms in the cascade from the initial scale down to the hadronisation scale. For a process with n shower branchings, there are thus in general $2^n n!$ histories in a CS dipole shower. Conventional antenna (and Lund-dipole) showers reduce this number by 2^n, as there is now only one term for each *pair* of colour-connected partons, but the scaling[1] still goes like $n!$. Multileg merging therefore to date remains impeded by the sheer number of possible histories. This is especially true for CS-style approaches while antenna-based ones (see, e.g., [39]) should be expected to exhibit a somewhat lower computational overhead, see table 1.

A promising alternative that could reduce the complexity even further is the *sector antenna* formalism [38, 43–46]. In the context of the early Vincia final-state shower, it was shown in [38] that an antenna-shower history can be made unique if instead of the conventional (henceforth called *global*) shower algorithm, a *sector* formulation of the antenna framework is used, in which only a single shower term is allowed to contribute to each specific $(n+1)$-parton phase-space point; all other potential contributions are vetoed. In this framework, each antenna splitting kernel must incorporate both the full soft and collinear limits of the respective phase-space sectors, in contrast to global antenna functions (such as those used in Ariadne and in earlier versions of Vincia) which smoothly partition the collinear $g \mapsto gg$ singularities between two neighbouring antennae. Although

[1]This can in principle be reduced to just n terms after the n-th branching by imposing a strictly Markovian ordering condition [12]. This is, however, likely to lead to undesired side effects [41, 42] and hence is not considered here.

JHEP07(2020)032

there remains an ambiguity in how to treat cases with multiple interfering Born processes,[2] the potential number of these remains small for all n.

In the sector framework, the sharing of the gluon-collinear limits is viewed as a discrete partitioning of phase space into two *sectors*, which in the collinear (DGLAP) limit correspond to $z < 1/2$ and $z > 1/2$ respectively; in this picture, the gluon with lower energy fraction z is *always* regarded as the emitted one. To generalise this *sector decomposition*, we use a notion of transverse momentum ($p_\perp^2$) which tends to zQ^2 in the collinear limits (with Q^2 the virtuality of the branching particle), so that it remains unique, well-defined, and exact outside the singular limits. It is then the gluon with the lowest $p_\perp^2$ which is regarded as the emitted one. Inside each sector only one antenna, which captures the full soft singularity of the respective phase-space sector and the full $z < \frac{1}{2}$ collinear singularities for quarks and gluons respectively, is allowed to contribute. This brings the number of possibilities to go from a given final state all the way back to the Born configuration down to one (modulo the question of interfering Born processes already alluded to), thus creating a unique shower history. Sector showers thus have great potential for speeding up matched and merged multijet calculations.

Helicity- and mass-dependent sector antenna functions for initial- and final-state radiation were already presented in [45, 46], although without a dedicated implementation in a shower algorithm. We here present a new implementation of a sector shower based on the Vincia antenna shower in the Pythia 8.3 framework,[3] including helicity dependence [47, 48], mass corrections [20], initial-state radiation [22, 49], resonance decays [50], and interleaved coherent QED branchings [51]. We define a full set of helicity-dependent sector antenna functions for initial- and final-state radiation entirely based on crossings and sums of global final-final antenna functions and validate the new shower against leading-order matrix elements, the global Vincia shower, Pythia 8.3, and experimental data.

The paper is structured as follows. After reviewing the foundations of the Vincia antenna shower in section 2, the sector-shower implementation is explained in detail in section 3. The implementation is validated in section 4 and we conclude and give an outlook on matching and merging applications in section 5.

2 The Vincia antenna shower

In this section, the main focus is on aspects that are common to both sector and global showers, including phase-space factorisations, kinematics maps (a.k.a. recoil strategies), and our choices of evolution variables. In addition to providing the starting point for our

[2] An example of two such interfering Born processes is $H \to gg$ and $H \to b\bar{b}$, which mix at the $H \to b\bar{b}g$ level, via the $gg \mapsto b\bar{b}g$ and $b\bar{b} \mapsto b\bar{b}g$ branchings, respectively. Note that we use the slightly different symbols $\to$ and $\mapsto$ to distinguish between hard processes and shower evolution steps.

[3] We note that, at the technical level, the major undertaking of integrating the Vincia shower model fully within Pythia 8.3 (where previously it was developed a stand-alone plugin to Pythia 8.2) was just recently completed. To simplify this work, only a minimal subset of Vincia — which only includes the global antenna-shower model — was incorporated into Pythia version 8.301. The sector shower, along with other features such as matrix-element corrections, will be included in forthcoming updates in the near future. The results we present in this paper are obtained with a development version that essentially represents Pythia 8.301 + sector shower.

JHEP07(2020)032

discussion of sector showers in section 3, the material in this section therefore also serves as an up-to-date summary of the current implementation of VINCIA's global showers in PYTHIA 8.3.

The starting point for antenna showers is the factorisation of QCD amplitudes in the soft-gluon limit,

$$|\mathcal{M}_{n+1}(1,\ldots,i,j,k,\ldots,n+1)|^2 \overset{g_j \text{ soft}}{\longrightarrow} a_{j/IK}(s_{ij},s_{jk})|\mathcal{M}_n(1,\ldots,I,K,\ldots,n)|^2\,, \tag{2.1}$$

where we adopt the convention of [49] and label parent or "pre-branching" partons (i.e., ones in the n-parton configuration) by capital letters and daughter or "post-branching" partons (i.e., ones in $(n+1)$-parton configuration) by lowercase letters. Furthermore, legs that are in the initial state will be denoted with letters from the beginning of the alphabet, a, b, while final-state legs are denoted with letters starting from i, j, k. Dimensionful invariants, regardless of whether the partons are incoming or outgoing, are cast in terms of dot products

$$s_{jk} \equiv 2p_j \cdot p_k\,, \tag{2.2}$$

while invariant masses for final-final (FF) and initial-initial (II) antennae are denoted as

$$m^2_{jk} = (p_j + p_k)^2 = m^2_j + m^2_k + s_{jk}\,, \tag{2.3}$$

and momentum transfers for initial-final (IF) and resonance-final (RF) antennae as

$$q^2_{ai} = (p_a - p_i)^2 = m^2_a + m^2_i - s_{ai}\,. \tag{2.4}$$

Energy-momentum conservation then implies:

$$\text{FF}: \qquad s_{IK} + m^2_I + m^2_K = s_{ij} + s_{ik} + s_{jk} + m^2_i + m^2_j + m^2_k \tag{2.5}$$

$$\text{RF \& IF}: \qquad s_{AK} - m^2_A - m^2_K = s_{aj} + s_{ak} - s_{jk} - m^2_a - m^2_j - m^2_k \tag{2.6}$$

$$\text{II}: \qquad s_{AB} + m^2_A + m^2_B = s_{ab} - s_{aj} - s_{jb} + m^2_a + m^2_j + m^2_b \tag{2.7}$$

and we define dimensionless invariants $y \in [0,1]$

$$y_{ij} = \frac{s_{ij}}{s} \qquad\qquad \mu^2_j = \frac{m^2_j}{s} \tag{2.8}$$

by scaling by the largest dynamical invariant,[4] s:

$$s = \begin{cases} s_{IK} & \text{FF} \\ s_{aj} + s_{ak} & \text{RF and IF} \\ s_{ab} & \text{II} \end{cases}\,. \tag{2.9}$$

It should be emphasised that parton j is always in the final state, as we do not consider emission into the initial state. In section 2.3 on branching kinematics below, we provide

[4] In [49, 50] the normalisation $s_{AK} + s_{jk}$ was used, which differs from eq. (2.9) only for the case of massive splittings. The alternative normalisation used here was chosen in order to have better uniformity of conventions.

JHEP07(2020)032

some translations between the explicitly Lorentz-invariant y variables defined here and the z or x energy fractions that are used in other shower formalisms and in PDFs. Note also that for the specific case of gluon emission, $m_j = 0$ and the masses of the pre-branching partons are the same as those of the post-branching ones, hence all the m^2 terms in eqs. (2.5) and (2.7) cancel. Those terms are only relevant for branching processes involving a change in the number of massive quarks.

Considering it a better approximation to use massless kinematics for incoming heavy-flavour legs than to assign them their nominal on-shell values, the current treatment of initial-state mass effects in Vincia is a compromise, with all initial-state legs forced to have massless kinematics. This is similar to the choice made in Pythia's default showers [52]. Nevertheless, to maintain maximum generality and to avoid needlessly obscuring the crossing relations between initial- and final-state antennae, we maintain a general language where possible, which at least in principle allows for massive initial-state partons. Mass terms for initial-state partons are therefore included in the antenna functions, cf. appendix A, and in our evolution and sector resolution variables, cf. sections 2.5 and 3.3. When interpreting these in the context of massless initial-state kinematics, we use the following conventions:

- Motivated by [53], we include mass terms for both quarks in initial-state $q_A \mapsto g_a q_j$ branchings, which, from the forward-evolution point of view, look like the gluon splitting $g_a \mapsto \bar{q}_A q_j$. We justify this choice by noting that this branching is a crossing $j \to a$ of the final-state gluon emission process $q_I \mapsto q_i g_j$, with gluon g_j being crossed to be identified with g_a, and therefore no massive parton is crossed into the initial state.

- Following the same logic, we include mass terms only for the final-state quark in quark conversion $g_A \mapsto q_a q_j$, which, from the forward-evolution point of view, looks like an initial-state quark is radiating a gluon. There, the massive quark q_i in the $g_I \mapsto \bar{q}_i q_j$ splitting had to be crossed into the initial state to become the massive quark q_a.

Although some initial-state mass effects might be implemented without major inconsistencies in the near future, a full-fledged and consistent treatment of massive initial-state legs depends upon the availability of massive PDFs and is outside the scope of this work. First steps towards a consistent inclusion of initial-state mass effects have been presented in [54], where an implementation in the Catani-Seymour dipole shower in Sherpa has been outlined. As antennae can effectively be thought of as the sum of two CS dipoles, a similar scheme might be implemented in the Vincia antenna shower.

2.1 Antenna functions

The antenna function $a_{j/IK}$ in eq. (2.1) acts as the splitting kernel of the coherent $2 \mapsto 3$ branching $IK \mapsto ijk$ and may formally be represented as

$$a_{j/IK}(s_{ij}, s_{jk}, m_i^2, m_j^2, m_k^2) = \frac{|\mathcal{M}_3(i,j,k)|^2}{|\mathcal{M}_2(I,K)|^2}\,. \tag{2.10}$$

JHEP07(2020)032

in the factorised form of the integral over the three-body matrix element

$$|\mathcal{M}_3(i,j,k)|^2 \mathrm{d}\Phi_3 = |\mathcal{M}_2(I,K)|^2 \mathrm{d}\Phi_2 \, a_{j/IK}(s_{ij}, s_{jk}, m_i^2, m_j^2, m_k^2) \frac{\mathrm{d}\Phi_3}{\mathrm{d}\Phi_2}. \tag{2.11}$$

By construction, it reproduces the full eikonal factor in the limit of gluon g_j becoming soft,

$$a_{j/IK}(s_{ij}, s_{jk}, m_i^2, 0, m_k^2) \overset{g_j \text{ soft}}{\longrightarrow} g_\mathrm{s}^2 \mathcal{C}_{j/IK} \left[\frac{2s_{ik}}{s_{ij}s_{jk}} - \frac{2m_i^2}{s_{ij}^2} - \frac{2m_k^2}{s_{jk}^2} \right], \tag{2.12}$$

with a process-dependent colour factor $\mathcal{C}_{j/IK}$. When g_j becomes (quasi-)collinear with a quark i, it reproduces the full (quasi-collinear) DGLAP splitting kernel $P(z)$ (or $P(z)/z$ for an initial-state parton),

$$a_{j/IK}(s_{ij}, s_{jk}, 0, 0, 0) \overset{i \| j}{\longrightarrow} g_\mathrm{s}^2 \mathcal{C}_{j/IK} \frac{P(z)}{s_{ij}}, \tag{2.13}$$

where z is the energy fraction taken by quark i.

For the branchings $g \mapsto gg$, however, the treatment of the collinear limits differs between the global and sector formalisms. For the global antenna functions, the DGLAP splitting kernel is partitioned onto two neighbouring (colour-adjacent) antennae, whose collinear limits are related by $z \leftrightarrow 1-z$ so that one of them incorporates the $1/(1-z)$ part of the DGLAP kernel while the other incorporates the $1/z$ part. For each value of z the full collinear singularity is only recovered after summing over these two terms. In the sector formalism, instead, the shared gluon-collinear singularity is fully incorporated into *both* of the neighbouring antenna functions; these are then supplemented by a phase-space veto such that the $z \leq 1/2$ part of the collinear boundary is covered by only one of them and the $z > 1/2$ part by the other. We return to this point in section 3.1.

By contrast, due to the lack of a soft singularity, gluon splittings are handled somewhat differently in the antenna framework. In gluon splittings, a clear distinction between the splitter and spectator is possible; therefore these could in principle be treated differently to gluon emissions. However, since energy and momentum conservation for on-shell partons requires at least a $2 \mapsto 3$ phase-space factorisation and since we want to generate a common (interleaved) sequence in which all types of evolution steps are evolved in a common resolution measure [52], we use the same phase-space and kinematics maps and ordering variables as for gluon emissions.[5]

To distinguish between global and sector antenna functions, and between which partons are initial- and which are final-state ones, we label them as follows:

$$a_{j/IK}^{\text{state,type}}(s_{ij}, s_{jk}, m_i^2, m_j^2, m_k^2), \tag{2.14}$$

with "state" $\in$ [II, IF, FF, RF], "type" $\in$ [glb,sct].

Note that since $2 \mapsto 3$ antenna functions have dimension -2 in natural units, we usually cast them in terms of a dimensionless function of the y and μ variables defined in

[5] We note that this choice differs from that of the previous version of Vincia [49], in which gluon emissions were evolved in $p_\perp$ while gluon splittings were evolved in virtuality.

JHEP07(2020)032

eq. (2.8) multiplied by a single (constant) dimensionful quantity,[6] cf. appendix A. Moreover, throughout we use colour- and coupling-stripped antenna functions $\bar{a}$,

$$a_{j/IK}(s_{ij}, s_{jk}, m_i^2, m_j^2, m_k^2) = 4\pi\,\alpha_s\,\mathcal{C}_{j/IK}\,\bar{a}_{j/IK}(y_{ij}, y_{jk}, \mu_i^2, \mu_j^2, \mu_k^2)\,. \tag{2.15}$$

Below, we will often let the mass dependence be implicit to avoid clutter. The colour factor $\mathcal{C}_{j/IK}$ is chosen in a convention in which both C_F and C_A tend to N_C in the large-N_C limit and where T_R is unity. In this convention, the ambiguity related to the colour factor of a qg antenna is explicitly a subleading-colour effect [11].

The complete set of antenna functions used in VINCIA are collected in appendix A, including both mass and helicity dependence.

2.2 Phase-space factorisation

The Lorentz-invariant $(n+1)$-particle phase space measure in four dimensions,

$$\mathrm{d}\Phi_{n+1}(p_a, p_b; p_1, \ldots, p_{n+1}) = (2\pi)^4\delta^{(4)}\left(p_a + p_b - \sum_{\ell=1}^{n+1} p_\ell\right)\prod_{\ell=1}^{n+1}\frac{\mathrm{d}^4 p_\ell}{(2\pi)^3}\delta(p_\ell^2 - m_\ell^2)\,, \tag{2.16}$$

exactly factorises into the $(n-m+1)$-particle phase-space measure $\mathrm{d}\Phi_{n-m+1}$ and the m-particle branching measure $\widetilde{\mathrm{d}\Phi}_m$,

$$\mathrm{d}\Phi_{n+1}(p_a, p_b; p_1, \ldots, p_j, \ldots, p_{n+1}) = \mathrm{d}\Phi_{n-m+1}(p_A, p_B; p_1, \ldots, p_{n+1})\widetilde{\mathrm{d}\Phi}_m\,. \tag{2.17}$$

FF branchings. The branching measure $\widetilde{\mathrm{d}\Phi}_m$ is proportional to the three-particle phase space measure $\mathrm{d}\Phi_3$ which can be factorised into the product of the two-particle pre-branching phase space and the FF antenna phase space measure,

$$\mathrm{d}\Phi_3(p_A, p_B; p_i, p_j, p_k) = \mathrm{d}\Phi_2(p_A, p_B; p_I, p_K)\mathrm{d}\Phi_{\mathrm{ant}}^{\mathrm{FF}}(p_i, p_j, p_k)\,. \tag{2.18}$$

In terms of the dimensionless invariants y_{ij}, y_{jk} and an angle ϕ between the branching plane and the parent dipole, the final-final antenna phase space can be written as [20]

$$\mathrm{d}\Phi_{\mathrm{ant}}^{\mathrm{FF}} = \frac{1}{16\pi^2} f_{\mathrm{Källén}}^{\mathrm{FF}} s_{IK}\Theta(\Gamma_{ijk})\mathrm{d}y_{ij}\mathrm{d}y_{jk}\frac{\mathrm{d}\phi}{2\pi}\,, \tag{2.19}$$

with the dimensionless three-body Gram determinant expressing boundaries of the physical phase space

$$\Gamma_{ijk} = y_{ij}y_{jk}y_{ik} - y_{jk}^2\mu_i^2 - y_{ik}^2\mu_j^2 - y_{ij}^2\mu_k^2 + 4\mu_i^2\mu_j^2\mu_k^2\,, \tag{2.20}$$

again scaled by the appropriate normalisation according to eq. (2.8). Note that for massless partons, the $\Theta(\Gamma)$ factor just reduces to the standard triangular phase space defined by $y_{ij} \geq 0$, $y_{jk} \geq 0$, $y_{ij} + y_{jk} \leq 1$, while for massive ones it defines a smaller region inside this hull, see, e.g. [20]. The volume of the two-particle phase-space of the parent antenna is taken into account by the Källén factor,

$$f_{\mathrm{Källén}}^{\mathrm{FF}} = \frac{s_{IK}}{\sqrt{\lambda(m_{IK}^2, m_I^2, m_K^2)}}\,, \tag{2.21}$$

[6] Specifically we normalise by s_{IK}^{-1}, s_{AK}^{-1} and s_{AB}^{-1} for the FF, IF(RF) and II antenna functions respectively.

JHEP07(2020)032

in terms of the Källén function

$$\lambda(a,b,c) = a^2 + b^2 + c^2 - 2(ab + ac + bc)\,. \tag{2.22}$$

Note that $f^{\rm FF}_{\rm Källén} = 1$ if either or both of partons I and K are massless, while when both are massive we have $f^{\rm FF}_{\rm Källén} > 1$. This does not present a problem for generating numerical overestimates in the shower; $f^{\rm FF}_{\rm Källén}$ is just an overall constant and can be factored out of the branching integrals.

RF branchings. The phase space factorisation for resonance-final branchings is largely unchanged relative to the final-final case [50]:

$$\mathrm{d}\Phi^{\rm RF}_{\rm ant} = \frac{1}{16\pi^2} f^{\rm RF}_{\rm Källén} \frac{s_{AK} + m_j^2 + m_k^2 - m_K^2}{(1-y_{jk})^3} \Theta(\Gamma_{ajk}) \mathrm{d}y_{aj} \mathrm{d}y_{jk} \frac{\mathrm{d}\phi}{2\pi}\,, \tag{2.23}$$

where now the Källén factor is given by:

$$f^{\rm RF}_{\rm Källén} = \frac{s_{AK} + m_j^2 + m_k^2 - m_K^2}{\sqrt{\lambda(m_A^2, m_{AK}^2, m_K^2)}}\,. \tag{2.24}$$

IF branchings. Keeping initial-state partons explicitly massless but allowing for massive final-state ones, the factorisation eq. (2.18) is replaced by the convolution [22, 53, 55]

$$\begin{aligned}\int \frac{\mathrm{d}x_a}{x_a}\Theta(1-x_a)\frac{\mathrm{d}x_B}{x_B}\Theta(1-x_B)\mathrm{d}\Phi_3(p_a,p_B;p_j,p_k,p_R) = \\ \int \frac{\mathrm{d}x_A}{x_A}\Theta(1-x_A)\frac{\mathrm{d}x_B}{x_B}\Theta(1-x_B)\mathrm{d}\Phi_2(p_A,p_B;p_K,p_R)\mathrm{d}\Phi^{\rm IF}_{\rm ant}(p_a,p_j,p_k)\,,\end{aligned} \tag{2.25}$$

where the IF antenna phase space measure can be written in terms of dimensionless invariants y_{aj}, y_{jk}, and the angle ϕ as [56]

$$\mathrm{d}\Phi^{\rm IF}_{\rm ant} = \frac{1}{16\pi^2}\frac{1}{1-y_{jk}} s_{AK}\Theta(\Gamma_{ajk})\mathrm{d}y_{aj}\mathrm{d}y_{jk}\frac{\mathrm{d}\phi}{2\pi}\,. \tag{2.26}$$

II branchings. With a massive emitted final-state parton, the branching measure is given by the convolution [22, 55]

$$\begin{aligned}\int \frac{\mathrm{d}x_a}{x_a}\Theta(1-x_a)\frac{\mathrm{d}x_b}{x_b}\Theta(1-x_b)\mathrm{d}\Phi_2(p_a,p_b;p_j,p_r) \\ = \int \frac{\mathrm{d}x_A}{x_A}\Theta(1-x_A)\frac{\mathrm{d}x_B}{x_B}\Theta(1-x_B)\mathrm{d}\Phi_1(p_A,p_B;p_R)\mathrm{d}\Phi^{\rm II}_{\rm ant}(p_a,p_j,p_b)\,,\end{aligned} \tag{2.27}$$

where again, the antenna phase space can be written in terms of dimensionless invariants y_{aj}, y_{jb}, and the angle ϕ as [56]

$$\mathrm{d}\Phi^{\rm II}_{\rm ant} = \frac{1}{16\pi^2}\frac{1}{1-y_{aj}-y_{jb}} s_{AB}\Theta(\Gamma_{ajb})\mathrm{d}y_{aj}\mathrm{d}y_{jb}\frac{\mathrm{d}\phi}{2\pi}\,. \tag{2.28}$$

JHEP07(2020)032

Common form. In general the antenna phase space can be written more compactly as

$$\mathrm{d}\Phi_{\mathrm{ant}} = \frac{1}{16\pi^2} F_\Phi \Theta(\Gamma_{ijk}) \mathrm{d}y_{ij} \mathrm{d}y_{jk} \frac{\mathrm{d}\phi}{2\pi}\,, \tag{2.29}$$

where we have introduced the (possibly dynamical) phase space factor F_Φ defined by

$$F_\Phi = \begin{cases} f^{\mathrm{FF}}_{\mathrm{Källén}} s_{IK} & \mathrm{FF} \\ f^{\mathrm{RF}}_{\mathrm{Källén}} \dfrac{s_{AK} + m_j^2 + m_k^2 - m_K^2}{(1-y_{jk})^3} & \mathrm{RF} \\ \dfrac{s_{AK}}{1-y_{jk}} & \mathrm{IF} \\ \dfrac{s_{AB}}{1-y_{aj}-y_{jb}} & \mathrm{II} \end{cases} . \tag{2.30}$$

2.3 Branching kinematics

In the antenna-shower formalism both of the parents I and K in an FF branching $IK \mapsto ijk$ act collectively as emitters, and the transverse recoil of parton j is shared between them. In IF or II branchings, this picture is changed slightly by the requirement that initial-state partons must be aligned with the beam axis. In these cases, the kinematic mappings are closer related to those used in the CS-type dipole picture [16]. In RF branchings, the on-shell momentum of the initial decaying resonance is fixed; in the resonance rest frame, the transverse recoil is shared between the two daughters directly involved in the branching, with all remaining decay products taking the longitudinal recoil required to preserve the decaying resonance's mass [50]. In all cases, all partons remain on-shell and all four-momenta are conserved at every step in the antenna shower evolution.

Although we derive all of our antenna functions for general masses of the involved partons, cf. appendix A, we construct the kinematics for massless initial-state partons in order to not contravene the assumptions of collinear PDF evolution. We do however let any emitted final-state quark, denoted by j, acquire a finite mass if required and refer the reader to section 2 for our conventions on massive quarks in initial-state gluon splittings. Should it become feasible to incorporate initial-state masses in the future, the kinematics maps below could be altered for instance by means of the kinematic maps in [19].

FF branchings. The post-branching momenta are constructed in the centre-of-mass frame of the parent antenna in two steps. In the first step, the momenta are defined in a way that is agnostic to the overall orientation of the parent partons. Choosing (arbitrarily) to align parton i with the z axis of this temporary coordinate system [20], we have:

$$p_i^\mu = (E_i, 0, 0, |\vec{p}_i|)\,, \tag{2.31}$$

$$p_j^\mu = (E_j, -|\vec{p}_j| \sin\theta_{ij}, 0, |\vec{p}_j| \cos\theta_{ij})\,, \tag{2.32}$$

$$p_k^\mu = (E_k, |\vec{p}_k| \sin\theta_{ik}, 0, |\vec{p}_k| \cos\theta_{ik})\,, \tag{2.33}$$

JHEP07(2020)032

with energies, E, (and energy fractions, x_E)

$$E_i = \frac{s_{ij}+s_{ik}+2m_i^2}{2m_{IK}}\,, \qquad E_j = \frac{s_{ij}+s_{jk}+2m_j^2}{2m_{IK}}\,, \qquad E_k = \frac{s_{ik}+s_{jk}+2m_k^2}{2m_{IK}}\,, \tag{2.34}$$

$$\Longrightarrow \quad x_{Ei} \equiv \frac{2E_i}{m_{IK}} = 1 - y_{jk} - \frac{\mu_k^2}{1+\mu_k^2}(2-y_{jk})\,, \qquad \text{same for } x_{Ek} \text{ with } i \leftrightarrow k\,, \tag{2.35}$$

and angles

$$\cos\theta_{ij} = \frac{2E_iE_j - s_{ij}}{2|\vec{p}_i||\vec{p}_j|}\,, \qquad \cos\theta_{ik} = \frac{2E_iE_k - s_{ik}}{2|\vec{p}_i||\vec{p}_k|}\,, \tag{2.36}$$

and the on-shell conditions for $|\vec{p}_{i,j,k}| = \sqrt{E_{i,j,k}^2 - m_{i,j,k}^2}$.

In the second step, the branching plane is rotated by an angle ϕ, uniformly sampled in $[0, 2\pi]$, in the x-y-plane, and by the angle ψ between the mother parton I and the daughter parton i. It is this latter rotation by ψ, which establishes the relative orientation of the post-branching partons with respect to the axis defined by the pre-branching ones. The choice of ψ is not unique outside of the collinear limits[7] and a few different options for the choice of ψ are implemented in VINCIA, see [11, 20]. The ambiguity in the choice of ψ can be expressed in terms of a single free parameter r, as

$$\cos\psi = \frac{2E_IE_i - 2c_im_i^2 - rs_{ij} - (1-c_k)s_{ik}}{2|\vec{p}_I||\vec{p}_i|}\,, \tag{2.37}$$

where the (kinematics-dependent) constants $c_{i,k}$ are determined by requiring on-shellness of all particles in the $\{I, K\} \mapsto \{i, j, k\}$ mapping, cf. [20, 43],

$$p_I = c_ip_i + rp_j + (1-c_k)p_k\,, \tag{2.38}$$

$$p_K = c_kp_k + (1-r)p_j + (1-c_i)p_i\,. \tag{2.39}$$

It is worth noting that in the limit $r \to 1$, these expressions imply that any transverse momentum carried by parton j is fully absorbed by parton i, hence we require that any sensible choice of r should tend to unity in the I-collinear limit. Similarly, r should vanish in the K-collinear limit. The phase space boundaries are given by $0 \leq y_{ij} \leq 1$ and $0 \leq y_{jk} \leq 1 - y_{ij}$ as well as the roots of the Gram determinant.

For massless partons, our default choice for r is:

$$r \;=\; \frac{s_{jk}}{s_{ij}+s_{jk}} \;\to\; \begin{cases} 1 & I\text{-collinear limit, } s_{ij} \ll s_{jk} \\ 0 & K\text{-collinear limit, } s_{jk} \ll s_{ij} \end{cases}\,, \tag{2.40}$$

which therefore obeys the requirement given above. For massive partons, the form is more involved, see [20].

[7] In the K-collinear limit, $\psi \to 0$ ensures that parton i recoils purely longitudinally along the direction of parton I, and similarly $\psi \to \pi - \theta_{ik}$ in the I-collinear limit.

JHEP07(2020)032

RF branchings. The post-branching kinematics are constructed according to the prescription in [50]. With the invariant mass of the resonance being kept fixed, the post-branching momenta are constructed in the resonance rest frame, such that

$$p_A^\mu = p_a^\mu = (m_A, 0, 0, 0) \,, \tag{2.41}$$

$$p_k^\mu = \left(E_k, 0, 0, \sqrt{E_j^2 - m_k^2}\right) \,, \tag{2.42}$$

$$p_j^\mu = \left(E_j, \sqrt{E_j^2 - m_j^2}\sin\theta_{jk}, 0, \sqrt{E_j^2 - m_j^2}\cos\theta_{jk}\right) \,, \tag{2.43}$$

$$p_{X'}^\mu = \left(m_A - E_k - E_j, -\sqrt{E_j^2 - m_j^2}\sin\theta_{jk}, 0, -\sqrt{E_j^2 - m_k^2} - \sqrt{E_j^2 - m_j^2}\cos\theta_{jk}\right) \,, \tag{2.44}$$

where

$$E_j = \frac{s_{aj}}{2m_a}, \quad E_k = \frac{s_{ak}}{2m_a}, \quad \cos\theta_{jk} = \frac{2E_bE_g - s_{jk}}{2\sqrt{(E_k^2 - m_k^2)(E_j^2 - m_j^2)}} \,, \tag{2.45}$$

and where $\{X'\}$ denote the remainder of the resonance decay system. The frame orientation is chosen such that the z-axis is defined along p_K. These momenta are rotated about the y-axis such that the set of recoilers are along -z, so that only j and k receive transverse recoil. The momenta are subsequently rotated by a flatly sampled azimuthal angle ϕ about the z-axis. Finally, the original orientation of K with respect to z is recovered, before boosting back to the lab frame.

The conservation of the invariant mass of the system of recoilers, $p_X = \sum_{i\in\{X\}} p_i$ is automatically ensured by imposing eq. (2.6) and the positivity of the Gram determinant Γ_{ajk} defined in eq. (2.20). (Note, however, that this statement no longer applies if only a subset of X are chosen to receive the longitudinal recoil.)

IF branchings. For massless initial-state partons, A and a, and general final-state ones, K, j, and k, and requiring that the beam axis defined by the pre-branching incoming parton A remains the same post-branching, the post-branching momenta in the A-K rest frame are given by [56]

$$p_a^\mu = \frac{1}{y_{AK}} p_A^\mu \,, \tag{2.46}$$

$$p_j^\mu = \frac{(y_{ak} + \mu_j^2 - \mu_k^2) + (y_{ak} - y_{aj})\mu_K^2 - y_{AK}y_{ak}}{y_{AK}}\, p_A^\mu + y_{aj}\, p_K^\mu + \sqrt{\Gamma_{ajk}}\, q_{\perp\max}^\mu \,, \tag{2.47}$$

$$p_k^\mu = \frac{(y_{aj} - \mu_j^2 + \mu_k^2) + (y_{aj} - y_{ak})\mu_K^2 - y_{AK}y_{aj}}{y_{AK}}\, p_A^\mu + y_{ak}\, p_K^\mu - \sqrt{\Gamma_{ajk}}\, q_{\perp\max}^\mu \,, \tag{2.48}$$

where $q_{\perp\max}^\mu$ is a space-like four-vector in the transverse direction, perpendicular to both p_A and p_K, with $q_{\perp\max}^2 = -(s_{aj} + s_{ak})$. The angle about the branching plane is uniformly distributed in $[0, 2\pi]$.

In the massless case, this reduces to the form given in [49]. The phase space boundaries are given by $0 \le y_{aj} \le 1$ and $0 \le y_{jk} \le \frac{1-x_A}{x_A} y_{AK} - \mu_j^2 - \mu_k^2 + \mu_K^2$ as well as the roots of the Gram determinant.

JHEP07(2020)032

In the a-collinear limit, we see trivially that

$$z_A \equiv \frac{x_A}{x_a} = y_{AK} = \frac{s_{AK}}{s_{aj}+s_{ak}}, \tag{2.49}$$

while in the k-(quasi)-collinear limit, we find

$$z_K = 1 - y_{aj} = y_{ak} = \frac{s_{ak}}{s_{aj}+s_{ak}}. \tag{2.50}$$

We note that in addition to the "local" map discussed here, there is also a "global" map implemented in VINCIA, where instead of the final-state parton, it is the initial-state parton which recoils transversely. Moreover, VINCIA offers the possibility to interpolate between these two maps by choosing one of the two maps according to some probability imposed by comparison to matrix elements [56]. While this procedure ties in better with the recoiler-emitter-agnostic antenna formalism, we leave a dedicated study regarding the effect of choosing either of the latter two alternatives in the sector shower to future studies and here only use the local map for IF branchings.

II branchings. In an initial-initial branching, both partons a and b are assumed to stay aligned with the beam momenta p_A and p_B, respectively and only the emitted final-state parton j acquires mass, the post-branching momenta can be constructed to be [56]:

$$p_a^\mu = \sqrt{\frac{1}{y_{AB}}\frac{1-y_{aj}}{1-y_{jb}}}\,p_A^\mu\,, \tag{2.51}$$

$$p_j^\mu = \sqrt{\frac{y_{jb}^2}{y_{AB}}\frac{1-y_{aj}}{1-y_{jb}}}\,p_A^\mu + \sqrt{\frac{y_{aj}^2}{y_{AB}}\frac{1-y_{jb}}{1-y_{aj}}}\,p_B^\mu + \sqrt{y_{aj}y_{jb}-\mu_j^2}\,q_{\perp\max}^\mu\,, \tag{2.52}$$

$$p_b^\mu = \sqrt{\frac{1}{y_{AB}}\frac{1-y_{jb}}{1-y_{aj}}}\,p_B^\mu\,, \tag{2.53}$$

$$p_r^\mu = p_a^\mu + p_b^\mu - p_j^\mu, \tag{2.54}$$

again with $q_{\perp\max}^\mu$ a space-like four-vector perpendicular to p_A and p_B with $q_{\perp\max}^2 = -s_{ab}$ and the angle about the branching plane is uniformly distributed in $[0, 2\pi]$. The transverse momentum obtained by the emitted parton j is compensated for by the recoiler r, which denotes the rest of the system. Again, in the massless case, this is identical to the recoil scheme used in [49]. The phase space boundaries are $0 \leq y_{aj} \leq \frac{1-x_Ax_B}{x_Ax_B}y_{AB} + \mu_j^2$ and $0 \leq y_{jk} \leq \frac{1-x_Ax_B}{x_Ax_B}y_{AB} + \mu_j^2$ as well as the roots of the Gram determinant.

The condition that partons a and b stay aligned with the beam axis leads to the identity

$$\frac{x_Ax_B}{x_ax_a} = \frac{s_{AB}}{s_{ab}} = y_{AB}\,. \tag{2.55}$$

We round off by noting that eq. (2.55) implies that the z fractions on sides A and B are related by

$$z_A = z_B\frac{1-y_{jb}}{1-y_{aj}}\,. \tag{2.56}$$

JHEP07(2020)032

2.4 No-branching probability

Like other parton showers, the VINCIA antenna shower is built around the no-branching probability, i.e., the probability that no branching occurs between two scales $Q_n^2 > Q_{n+1}^2$, given by [22, 49]

$$\Pi_n(Q_n^2, Q_{n+1}^2) = \exp\left(-\sum_{i\in\{n\mapsto n+1\}} \mathcal{A}_i\left(Q_n^2, Q_{n+1}^2\right)\right), \tag{2.57}$$

where

$$\mathcal{A}_i(Q_n^2, Q_{n+1}^2) = \int_{Q_{n+1}^2}^{Q_n^2} 4\pi\alpha_s(Q^2)\mathcal{C}\bar{a}_i(Q^2,\zeta)R_f \mathrm{d}\Phi_{\mathrm{ant}}\,, \tag{2.58}$$

with a PDF ratio for each initial-state leg,

$$R_f = \begin{cases} 1 & \text{FF \& RF} \\ \dfrac{f_a(x_a,Q^2)}{f_A(x_A,Q^2)} & \text{IF} \\ \dfrac{f_a(x_a,Q^2)}{f_A(x_A,Q^2)}\dfrac{f_b(x_b,Q^2)}{f_B(x_B,Q^2)} & \text{II} \end{cases}. \tag{2.59}$$

When starting the evolution at a scale Q_n^2, the probability for a parton to branch at scale Q_{n+1}^2 is therefore given by

$$P_{\mathrm{branch}}(Q_n^2, Q_{n+1}^2) = \frac{\mathrm{d}\Pi_n(Q_n^2, Q_{n+1}^2)}{\mathrm{d}\log(Q_{n+1}^2)}\,. \tag{2.60}$$

Branchings are then generated by solving eq. (2.60) by means of the veto algorithm [57], by overestimating the antenna function by a larger trial function $\bar{a}_{\mathrm{trial}}$, for which a trial integral with an ansatz like $\hat{R}_f = \text{const}$ or $\hat{R}_f \propto (x_A x_B)/(x_a x_b)$ can be analytically solved to give simple expressions. By accepting each trial branching with a probability

$$P_{\mathrm{accept}} = \frac{\alpha_s}{\hat{\alpha}_s}\frac{\mathcal{C}}{\hat{\mathcal{C}}}\frac{R_f}{\hat{R}_f}\frac{\bar{a}}{\bar{a}_{\mathrm{trial}}}\,, \tag{2.61}$$

the full integral is recovered post-facto. The additional ratios

$$\frac{\alpha_s}{\hat{\alpha}_s}\,, \quad \frac{\mathcal{C}}{\hat{\mathcal{C}}}\,, \quad \text{and} \quad \frac{R_f}{\hat{R}_f}\,, \tag{2.62}$$

where hats denote trial quantities, take into account that nominally larger values than the physical ones may be used during the trial generation, cf. section 3.2.

2.5 Evolution variables

Both gluon emissions and gluon splittings are evolved in a scaled notion of off-shellness based on the ARIADNE transverse momentum, which we denote by $p_\perp^2$. An alternative (hypothetical) treatment permitting different definitions for the evolution measures of emissions and splittings (see, e.g., [49]) would in principle allow to use sector resolution variables, cf. section 3.3, as evolution variables for the sector shower. However in this case one would also have to address how this would impact the interleaving of the shower branchings; we therefore leave a dedicated study regarding this possibility to forthcoming work.

JHEP07(2020)032

FF branchings. We evolve in

$$p^2_{\perp,\mathrm{FF}} = \frac{(m^2_{ij}-m^2_I)(m^2_{jk}-m^2_K)}{s_{IK}} = \begin{cases} \dfrac{s_{ij}s_{jk}}{s_{IK}} & g\text{-emission} \\ \dfrac{m^2_{ij}(s_{jk}+m^2_j)}{s_{IK}} & g_I\text{-splitting} \end{cases}, \tag{2.63}$$

For gluon emissions, this evolution variable is identical to the ARIADNE $p^2_\perp$. For gluon splittings, the previous version of VINCIA [49] used the virtuality of the splitting gluon, m^2_{ij}, as the evolution variable, while we have here instead made the choice to evolve all branchings in a common measure of $p^2_\perp$ which, unlike m^2_{ij}, vanishes in the limit that the remaining jk colour dipole becomes unresolved.

RF branchings. Crossing partons I and i into the initial state to become partons A and a, we define the evolution variable for RF branchings as:

$$p^2_{\perp,\mathrm{RF}} = \frac{(m^2_A-q^2_{aj})(m^2_{jk}-m^2_K)}{s_{aj}+s_{ak}} = \begin{cases} \dfrac{s_{aj}s_{jk}}{s_{aj}+s_{ak}} & g\text{-emission} \\ \dfrac{m^2_{jk}(s_{aj}-m^2_j)}{s_{aj}+s_{ak}} & g_K\text{-splitting} \end{cases}. \tag{2.64}$$

IF branchings. In principle, the evolution variable for IF branchings is chosen to be the same as for RF ones, but with the additional cases of initial-state splittings and gluon conversions that are not present for RF branchings:

$$p^2_{\perp,\mathrm{IF}} = \frac{(m^2_A-q^2_{aj})(m^2_{jk}-m^2_K)}{s_{aj}+s_{ak}} = \begin{cases} \dfrac{s_{aj}s_{jk}}{s_{aj}+s_{ak}} & g\text{-emission} \\ \dfrac{m^2_{jk}(s_{aj}-m^2_j)}{s_{aj}+s_{ak}} & g_K\text{-splitting} \\ \dfrac{(s_{aj}-m^2_j)(s_{jk}+m^2_j)}{s_{aj}+s_{ak}} & q_A\text{-conversion} \\ \dfrac{s_{aj}(s_{jk}+m^2_j)}{s_{aj}+s_{ak}} & g_A\text{-splitting} \end{cases}, \tag{2.65}$$

where initial-state partons are treated as massless, $m_A = m_a = 0$, except for the case of a initial-state gluon splittings, where $m_a = m_j$, as alluded to above.

II branchings. We take into account that both parents, A and B are in the initial state, so that our choice of the evolution variable reads:

$$p^2_{\perp,\mathrm{II}} = \frac{(m^2_A-q^2_{aj})(m^2_B-q^2_{jb})}{s_{ab}} = \begin{cases} \dfrac{s_{aj}s_{jb}}{s_{ab}} & g\text{-emission} \\ \dfrac{(s_{aj}-m^2_j)(s_{jb}-m^2_j)}{s_{ab}} & q_A\text{-conversion} \\ \dfrac{s_{aj}(s_{jb}-m^2_j)}{s_{ab}} & g_A\text{-splitting} \end{cases}, \tag{2.66}$$

JHEP07(2020)032

with the same convention for massive partons in the initial state as in the IF case; i.e. both parton A and B are treated as massless, $m_A = m_B = 0$, except for splittings of initial-state gluons g_a or g_b, where $m_a = m_j$ or $m_b = m_j$, respectively.

3 The sector shower algorithm

In this section, we discuss the differences between the global antenna-shower implementation in Pythia 8.3, and the sector-shower model we have developed, including all aspects that are relevant for a full implementation of the sector-shower algorithm.

In section 3.1 we derive the physical sector antenna functions from the corresponding global ones. We then describe the overestimates used for trial generation and the components of the trial-generation algorithm in section 3.2. We close by reviewing our choice for the resolution variable used to discriminate between phase-space sectors in section 3.3.

3.1 Sector antenna functions

In conventional (global) showers, each radiation kernel is allowed to radiate into any phase-point that is kinematically accessible to it, subject only to the following requirements: (i) strong ordering (the evolution variable should be below the scale of the previous branching), (ii) perturbativity (it should be resolved with respect to the IR cutoff of the shower), and, potentially, (iii) that it should pass any angular-ordering or equivalent vetoes that are imposed in the given algorithm. Consequently, every phase-space point can receive contributions from multiple different parton-shower histories, which must be summed over to compute the "shower weight" for the phase-space point in question.

In the sector approach however, the phase space is divided into distinct sectors in each of which only a single antenna is allowed to contribute, cf. figures 3 and 4. In collinear limits corresponding to $g \mapsto gg$ or $g \mapsto q\bar{q}$ branchings, a single sector antenna function must therefore contain the same collinear singularity structure as the sum of two neighbouring global ones.

For $g \mapsto q\bar{q}$ branchings, there is no overlap with any soft singularity to worry about as far as the antenna functions are concerned,[8] and the sector antenna function can simply be taken to be twice the global one.

For $g \mapsto gg$ branchings, global antenna functions only contain explicit poles in $1/(1-z)$, not in $1/z$. However, since the collinear limits of neighbouring antennae are related by $z \leftrightarrow 1-z$, the full DGLAP kernel (which is symmetric under $z \leftrightarrow 1-z$),

$$P_{g\mapsto gg}(z) = \frac{(1-z(1-z))^2}{z(1-z)} = \frac{z}{1-z} + \frac{1-z}{z} + z(1-z)\,, \tag{3.1}$$

is recovered after summing over the two contributions,[9] see e.g. [49]. A sector antenna function on the other hand, must contain the full pole structure on its own, at least within

[8] There is a subtlety concerning how to choose the boundary between $g \mapsto q\bar{q}$ and $q \mapsto qg$ sectors however, which we return to below.

[9] One consequence of this is that global antenna functions are only unique up to terms that are antisymmetric under $z \leftrightarrow 1-z$, with e.g. Ariadne and GGG making different choices, while the global shower in Vincia allows a user-defined choice; see [12, 48].

JHEP07(2020)032

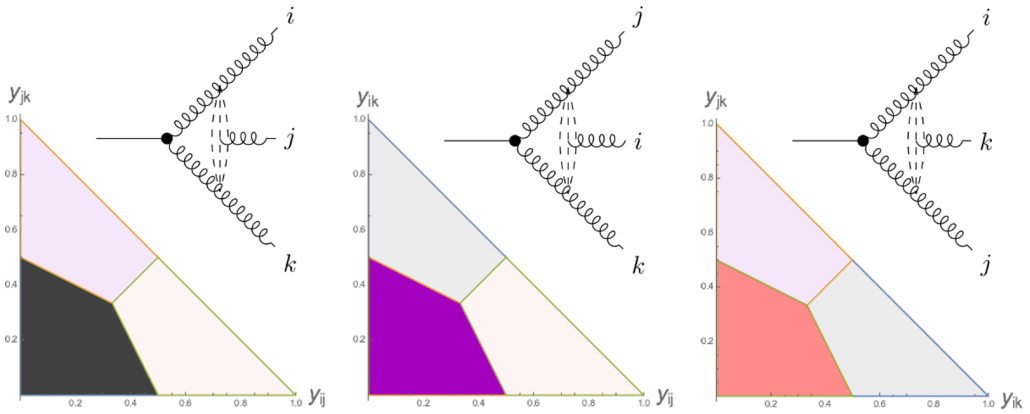


Figure 1. Illustration of the three different sectors of a $gg \mapsto ggg$ antenna, corresponding to the emission of gluon g_j (black), gluon g_i (blue), or gluon g_k (red). The sectors are shown in the respective antenna phase spaces, with the axes labelled such that the associated soft singularity is always located in the bottom left corner. Note that $z_i = 1 - y_{jk}$ in any i-collinear limit, and similarly for the j- and k-collinear limits.

the phase-space region the given sector is meant to cover. As we will discuss below, this corresponds to $z < \frac{1}{2}$ in the collinear limit since, in the sector context it is always the softer of the two gluons that is perceived of as the "emitted" one.

To illustrate this situation, take for instance the decay process $H \to gg$ followed by an FF branching process $gg \mapsto ggg$. There are three distinct shower histories contributing to the $H \to ggg$ final state (or six, in CS-style dipole approaches), illustrated by the three diagrams shown above the phase-space triangles in figure 1. In a global shower, each of the corresponding antenna functions (or, equivalently, CS dipole functions) radiates over all of the available phase-space region $y_{ij} + y_{jk} \leq 1$, and for each value of z the full collinear singularity involves an explicit sum over the antennae (or dipoles) which share the given collinear pair.

From the sector point of view, the phase space for the same $H \to ggg$ final state is regarded as composed of three distinct sectors, in which either g_i, g_j, or g_k, is considered as the emitted gluon, respectively. These sectors are illustrated by the differently shaded regions in the phase-space triangles shown in figure 1, using the $p^2_{\perp,\mathrm{FF}}$ variable defined in eq. (2.63) as the sector resolution criterion (discussed further in section 3.3). The phase-space sector which corresponds to, and is covered by, the antenna clustering shown above each triangle is shown with full shading, while the other two sectors are shown with partial shading.

Focusing on the left-hand plot, for instance, we see that the sector representing emission of gluon j (shown with full black shading) contains the entire singularity associated with gluon j becoming soft ($y_{ij} \to 0$ and $y_{jk} \to 0$) as well as the $z < \frac{1}{2}$ parts of the ij and jk collinear singularities ($y_{jk} = 1 - z_i > \frac{1}{2}$ and $y_{ij} = 1 - z_k > \frac{1}{2}$ respectively). In our approach, we nominally include the entire $(n+1)$-parton singularity structure in all three antenna functions, but only one of them is allowed to contribute in each sector.

Similar arguments apply to the gluon-collinear singularities in antennae with quark-gluon or gluon-antiquark parents. In such antennae, the global antenna functions already

JHEP07(2020)032

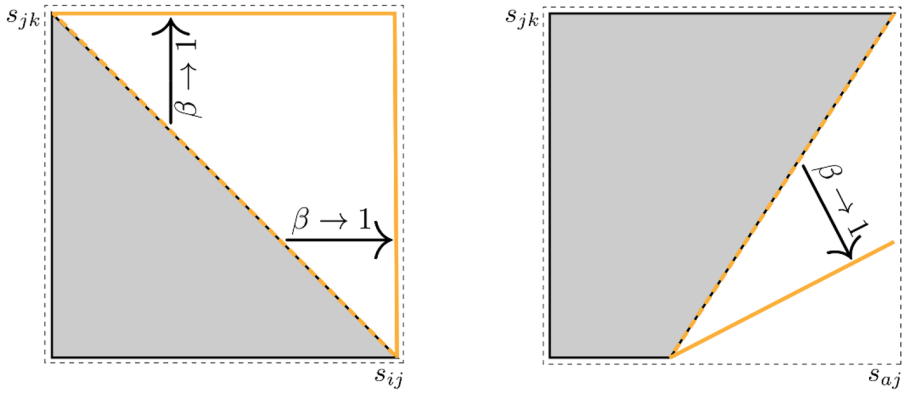


Figure 2. Illustration of the effect of the sector damp parameter β. For $\beta \to 1$, the unphysical singularity, marked in orange, is pushed away from the hard phase-space boundary $s_{ij} + s_{jk} = s_{IK}$ in the final-final phase-space (*left*) and $s_{jk} = s_{aj} - s_{AK}$ in the initial-final phase-space (*right*).

contain the full quark-collinear singularities, but again the gluon-collinear side only contains the $1/(1-z)$ poles, which must be summed together with $1/z$ ones from the neighbouring antenna to reproduce the full DGLAP kernel, phase-space point by phase-space point. In the sector approach, the full DGLAP limit is incorporated into both of the neighbouring antennae, but only the one corresponding to emission of the softer of the two gluons is allowed to contribute in each phase-space point. In the collinear limit, this is again equivalent to imposing a boundary at $z = \frac{1}{2}$ between the two histories.

We start from the global VINCIA FF antenna functions, cf. appendix A; global RF, IF, and II antenna functions are derived from these by crossing symmetry. Among these crossings are, formally, ones representing crossings of the emitted parton, j, into the initial state. Antennae corresponding to such crossings (emission into the initial state) are not part of the current VINCIA shower framework; instead the collinear singularities associated with these crossings are added onto the initial-state legs of the II and IF antennae, such that already in the global case the initial-state legs contain the full gluon-collinear singularities [22, 49].

To derive sector antenna functions based on global ones, we distinguish between gluon emissions and gluon splittings. As mentioned above, for the latter, simply a factor of two has to be included if the splitting happened in the final state, to account for the fact that the gluon is part of two global antennae. For gluon emissions, we include the missing gluon-collinear parts by symmetrising over each final-state gluon-gluon pair. For instance, the eikonal component of the global antenna function for $g_I g_K \mapsto g_i g_j g_k$, becomes

$$\text{Global: } \frac{2s_{ik}}{s_{ij}s_{jk}} \to \text{ Sector: } \frac{2s_{ik}}{s_{ij}s_{jk}} + \frac{2s_{ij}}{s_{jk}(s_{ik}+\beta s_{jk})} + \frac{2s_{jk}}{s_{ij}(s_{ik}+\beta s_{ij})}\,, \tag{3.2}$$

where the soft eikonal remains unchanged and the two additional terms correspond to the "missing" collinear parts in the ij- and jk-collinear limits. The auxiliary "sector damp" parameter $\beta \in [0,1]$, we have introduced here, allows to push the singularities for $s_{ik} \to 0$, that would result from an exact symmetrisation, away from the phase-space boundaries, cf. figure 2. Although this is not necessary for the construction of sector

JHEP07(2020)032

antenna functions, it ensures numerical stability in phase space configurations where the sector boundaries become close to the phase space boundary $s_{ik} = 0$. Moreover, as will be alluded to in section 3.2, it permits to generate trials over all of phase space with an a-posteriori imposed sector veto. Hence, our default choice is $\beta = 1$, but it can be changeed with `Vincia:sectorDamp`. We note that choosing $\beta \neq 0$ is not in disagreement with leading-logarithmic resummation as long as the sector resolution criterion reflects the singularity structure of the QCD matrix element, a point we return to in detail in section 3.3. For the example eq. (3.2) considered here, this implies that phase space points for which $s_{ik} \to 0$ always have to be assigned to the antenna where k is the emitted parton. This situation is shown in figure 1 for a simple three-parton configuration, where the radiation phase space is covered by three antennae, each of which is constrained to a unique region.

A further subtlety concerns the treatment of terms in the global antenna functions that are antisymmetric under $z \leftrightarrow 1 - z$. As mentioned previously, the contributions of such terms to the collinear limits cancel in the sum over two neighbouring global antennae. Outside of those limits, however, nonsingular differences can remain. In VINCIA, this ambiguity is parameterised by a parameter called the octet-partitioning parameter[10] α with default choice $\alpha = 0$, cf. appendix A. To ensure positivity of the sector antenna functions over all of their respective phase spaces, our sector antennae are based on global ones with $\alpha = 1$.

Below, we define our set of sector antenna functions derived from the above principles. For FF branchings, there are only slight differences with respect to the earlier sector-shower implementation of [38] while the IF, II, and RF aspects are new. The construction principles are applied to both helicity-averaged and, if the helicities of two collinear gluons match, helicity-dependent antenna functions — and helicity-dependence is implicitly understood. We discuss validating the collinear limits of sector antenna functions in appendix B.

FF branchings. For final-final gluon emission, we symmetrise over each gluon pair,

$$\begin{aligned}\bar{a}^{\mathrm{FF,sct}}_{g_j-\mathrm{emit}}(y_{ij}, y_{jk}) = {} & \bar{a}^{\mathrm{FF,gl}}_{g_j-\mathrm{emit}}(y_{ij}, y_{jk}) \\ & + \delta_{Ig}\delta_{h_I h_j}\bar{a}^{\mathrm{FF,gl}}_{g_i-\mathrm{emit}}(y_{ij}, 1 - y_{jk} - (1-\beta)y_{ij}) \\ & + \delta_{Kg}\delta_{h_K h_j}\bar{a}^{\mathrm{FF,gl}}_{g_k-\mathrm{emit}}(1 - y_{ij} - (1-\beta)y_{jk}, y_{jk}),\end{aligned} \tag{3.3}$$

with δ_{Ig} and δ_{Kg} being one if I or K is a gluon, respectively, and zero otherwise. Similarly, $\delta_{h_I h_j}$ and $\delta_{h_K h_j}$ are one if the helicity h_I or h_K of the parent gluon matches the one of the emitted gluon h_i. For $\beta \to 1$, the symmetrised invariants reduce to $1 - y_{jk}$ and $1 - y_{ij}$ for $i \leftrightarrow j$ and $j \leftrightarrow k$, respectively, ensuring that the additional, unphysical i-k-singularity is tamed.

For gluon splittings, we multiply by a factor of two,

$$\bar{a}^{\mathrm{FF,sct}}_{g-\mathrm{split}}(y_{ij}, y_{jk}) = 2\bar{a}^{\mathrm{FF,gl}}_{g-\mathrm{split}}(y_{ij}, y_{jk}) \tag{3.4}$$

where the gluon may either be g_I or g_K.

[10] It can be set with `Vincia:octetPartitioning`.

JHEP07(2020)032

RF branchings. Although in principle the antenna shower formalism can handle resonances of any spin and colour representation, so far only the antenna functions for spin-$\frac{1}{2}$, resonances in the fundamental representation (e.g. top quarks) have been included.[11] Thus in this case, the sector antenna functions for RF branchings are essentially unchanged with respect to the IF ones, discussed below, with the only restriction that the resonance cannot "backwards-evolve" (i.e., $a = A$).

IF branchings. In initial-final antennae, only final-state legs have to be sectorised, as there is no radiation into the initial state. For gluon emissions, we therefore have

$$\begin{aligned}\bar{a}^{\mathrm{IF,sct}}_{g_j-\mathrm{emit}}(y_{aj}, y_{jk}) = \bar{a}^{\mathrm{IF,gl}}_{g_j-\mathrm{emit}}(y_{aj}, y_{jk}) \\ + \delta_{Kg}\delta_{h_K h_j}\bar{a}^{\mathrm{IF,gl}}_{g_k-\mathrm{emit}}(1 - y_{aj} + \beta y_{jk}, y_{jk}),\end{aligned} \tag{3.5}$$

where again the additional, unphysical a-k-singularity resulting from the $j \leftrightarrow k$ symmetrisation is tamed for $\beta \to 1$.

For gluon splittings we include a factor of two only if the gluon was in the final state,

$$\bar{a}^{\mathrm{IF,sct}}_{g_K-\mathrm{split}}(y_{aj}, y_{jk}) = 2\bar{a}^{\mathrm{IF,gl}}_{g_K-\mathrm{split}}(y_{aj}, y_{jk}), \tag{3.6}$$

and use global antenna functions for initial-state gluon splittings and quark conversions,

$$\bar{a}^{\mathrm{IF,sct}}_{g_A-\mathrm{split}}(y_{aj}, y_{jk}) = \bar{a}^{\mathrm{IF,gl}}_{g_A-\mathrm{split}}(y_{aj}, y_{jk}), \tag{3.7}$$

$$\bar{a}^{\mathrm{IF,sct}}_{g_A-\mathrm{conv}}(y_{aj}, y_{jk}) = \bar{a}^{\mathrm{IF,gl}}_{g_A-\mathrm{conv}}(y_{aj}, y_{jk}). \tag{3.8}$$

II branchings. As already discussed, by construction initial-state legs are already sectorised, and so we can readily use VINCIA's global antenna functions in the sector shower without modifications,

$$\bar{a}^{\mathrm{II,sct}}(y_{aj}, y_{jb}) = \bar{a}^{\mathrm{II,gl}}(y_{aj}, y_{jb}). \tag{3.9}$$

3.2 Trial generation

Compared to the global case, sector antenna functions contain additional terms to incorporate the full gluon-collinear singularity in each sector. In the context of trial functions for gluon emissions, we note that the ij collinear limit of the eikonal overestimate used in VINCIA's global showers $\propto 1/p_\perp^2$ is $1/(Q_{ij}^2(1-z_i))$. This is sufficient to overestimate global antenna functions, since the $1/z_i$ part of the I-collinear singularity are contained in the neighbouring antenna, with the ij collinear limits of the two related by $z_i \leftrightarrow 1 - z_i$. For a sector shower, however, the eikonal $1/p_\perp^2$ does not overestimate the $1/z$ parts of the physical sector antenna functions. The sector shower therefore requires further trial generators for the collinear parts of the antennae, as we describe below.

We start by rewriting the antenna integral eq. (2.58) in the no-branching probability, eq. (2.57), in terms of dimensionless "shower variables", $x_\perp$ and ζ,

$$\mathcal{A}(x_{\perp 1}, x_{\perp 2}) = \frac{1}{16\pi^2}\int_{x_{\perp 2}}^{x_{\perp 1}} 4\pi\alpha_s\, \mathcal{C}\, \bar{a}(x_\perp, \zeta)\, R_f\, F_\Phi\, |J(x_\perp, \zeta)|\, \mathrm{d}x_\perp \mathrm{d}\zeta. \tag{3.10}$$

[11] Although it should be straightforward to include these, we postpone such a study of such hypothetical coloured scalar (e.g. gluinos) or vector (e.g. squark) resonances to future work.

JHEP07(2020)032

with the (dimensionless) variable $x_\perp$ defined by

$$x_\perp = \frac{p_\perp^2}{s}\,, \tag{3.11}$$

where $p_\perp^2$ is the evolution variable as defined in eqs. (2.63) to (2.66), and a complementary phase-space variable ζ which must be chosen such that the mapping

$$(y_{ij}, y_{jk}) \to (x_\perp, \zeta) \tag{3.12}$$

is one-to-one. The factor F_Φ is the phase-space factor defined in eq. (2.30) and $|J(x_\perp, \zeta)|$ denotes the Jacobian associated with the above variable transformation. As long as the Jacobian is properly accounted for, the choice of ζ does not affect physical observables. Below, this freedom in choosing ζ is exploited to produce relatively simple expressions for the trial integrals and to optimise the phase-space sampling for trial branchings.

From a given starting scale, $x_{\perp 1}$, the next branching scale in the downwards evolution is found by solving

$$R = \exp(-\mathcal{A}(x_{\perp 1}, x_{\perp 2})) \tag{3.13}$$

for $x_{\perp 2}$ with a uniformly distributed random number $R \in [0, 1]$. In general, this is not feasible analytically, hence we work instead with simple overestimates of the integrand in eq. (3.10) and use the Sudakov veto algorithm to ensure that the correct integral is recovered post-facto. That is, we replace physical antenna functions $\bar{a}$ by simpler trial functions $\bar{a}_{\text{trial}}$ that overestimate the physical antenna function everywhere in phase space and overestimate the PDF ratio R_f by

$$\hat{R}_f = \left(\frac{x_A}{x_a}\frac{x_B}{x_b}\right)^\gamma \frac{f_a(x_A, p^2_{\perp\min}) f_b(x_B, p^2_{\perp\min})}{f_A(x_A, p^2_{\perp\min}) f_B(x_B, p^2_{\perp\min})} =: \left(\frac{x_A}{x_a}\frac{x_B}{x_b}\right)^\gamma \tilde{R}_f, \tag{3.14}$$

with $p^2_{\perp\min}$ the minimal scale for the current trial and $\gamma \in \{0, 1\}$. The additional momentum-fraction ratio takes into account that for many cases (in particular for gluon and sea-quark distributions), an assumption that the PDFs fall off as some power of $1/x$ for higher x is a reasonable starting approximation. Where possible (except for valence-type flavours), our default choice is to use $\gamma = 1$.

We further assume α_s depends only on $p_\perp^2$ and define the one-loop running coupling

$$\hat{\alpha}_s(x_\perp) = \frac{1}{b_0 \log\left(\frac{x_\perp}{x_\Lambda}\right)}, \tag{3.15}$$

where

$$b_0 = \frac{33 - 2n_f}{12\pi} \quad \text{and} \quad \frac{x_\perp}{x_\Lambda} = \frac{k_R p_\perp^2}{\Lambda^2_{\text{QCD}}} \tag{3.16}$$

with an arbitrary scaling factor k_R to adjust the effective renormalisation scale. We note that two-loop running of the physical coupling is also implemented, via a simple modification of the veto probability, as done in [12].

JHEP07(2020)032

We overestimate all singular parts of the antenna functions individually and let different trials compete for the highest branching scale. A branching is then accepted with probability

$$P^{\text{sct}}_{\text{accept}} = \frac{\alpha_s}{\hat{\alpha}_s}\frac{\mathcal{C}}{\hat{\mathcal{C}}}\frac{R_f}{\hat{R}_f}\frac{\bar{a}}{\sum_{\text{trials}}\bar{a}_{\text{trial}}}, \tag{3.17}$$

Currently, we also utilise the veto algorithm to restrict branchings to the respective phase-space sectors, i.e., we generate trial branchings over all of phase space and accept only the one with minimal sector resolution variable of the post-branching configuration.

After accepting a branching, the complementary phase-space variable is generated by solving

$$R_\zeta = \frac{I_\zeta(\zeta_{\min}, \zeta)}{I_\zeta(\zeta_{\min}, \zeta_{\max})} \tag{3.18}$$

for ζ with a second uniformly distributed random number $R_\zeta \in [0, 1]$. Here, I_ζ denotes the integral over the ζ-dependence in $\mathcal{A}$, which is carried out over a larger region than the physically allowed phase space with simpler ζ boundaries. If this generates a value outside the physical boundaries,

$$\zeta < \zeta_{\min}(x_\perp) \quad \vee \quad \zeta > \zeta_{\max}(x_\perp), \tag{3.19}$$

the trial is vetoed and a new branching is generated with starting scale $x_\perp$.

To solve eq. (3.13) numerically, we rewrite the trial integral eq. (3.10) in terms of the function

$$\bar{\mathcal{A}}(x_{\perp 1}, x_{\perp 2}, \chi(\zeta)) = \int_{x_{\perp 2}}^{x_{\perp 1}} \int_{\zeta_{\min}(x_\perp)}^{\zeta_{\max}(x_\perp)} \alpha_s(x_\perp)\chi(\zeta)\mathrm{d}\zeta\frac{\mathrm{d}x_\perp}{x_\perp}. \tag{3.20}$$

for $\chi(\zeta)$ a function of ζ. For a one-loop running coupling, the solution to eq. (3.13) is then given by

$$x_{\perp 2} = x_\Lambda \left(\frac{x_{\perp 1}}{x_\Lambda}\right)^{R^{\frac{4\pi b_0}{\mathcal{C}\hat{R}_f I(\zeta_{\min},\zeta_{\max})}}}, \tag{3.21}$$

while for a constant trial coupling, it is given by

$$x_{\perp 2} = x_{\perp 1} R^{\frac{4\pi}{\bar{\alpha}_s \mathcal{C}\hat{R}_f I(\zeta_{\min},\zeta_{\max})}}, \tag{3.22}$$

or alternatively

$$x_{\perp 2} = \exp\left(-\sqrt{\log^2(x_{\perp 1}) - \frac{4\pi}{f^{\text{FF}}_{\text{Källén}}\alpha_s\mathcal{C}}\log(R)}\right). \tag{3.23}$$

if the zeta-integral can be evaluated to

$$I_{\log}(x_\perp) = \log\left(\frac{1}{x_\perp}\right). \tag{3.24}$$

Explicit expressions for eq. (3.20) are collected in appendix C.

JHEP07(2020)032

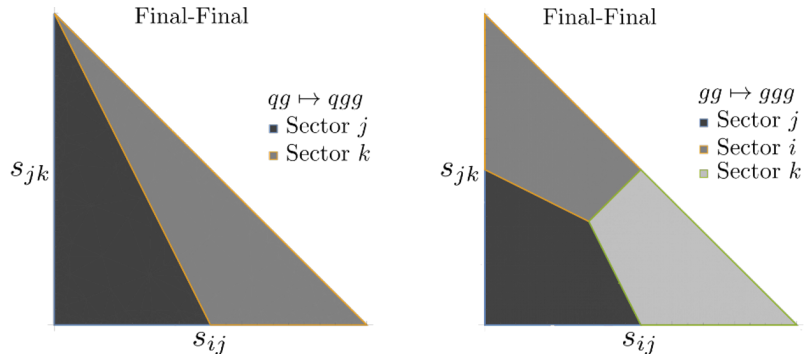


Figure 3. Illustration of the sectors in the antenna phase spaces for a final-final $qg \mapsto qgg$ antenna (*left*) and a final-final $gg \mapsto ggg$ antenna (*right*), with the sector resolution according to eq. (3.25).

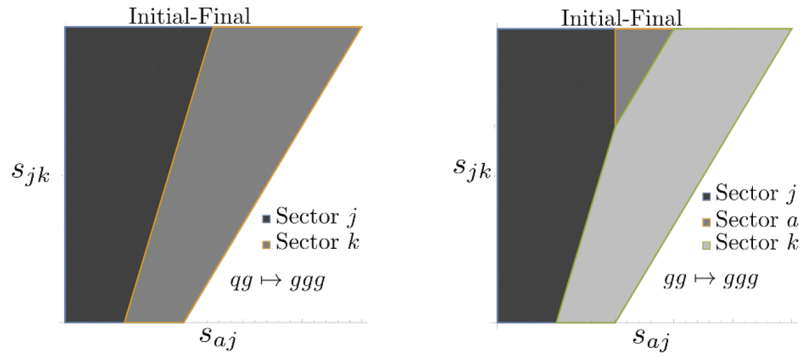


Figure 4. Illustration of the sectors in the antenna phase spaces for an initial-final $qg \mapsto qgg$ antenna (*left*) and an initial-final $gg \mapsto ggg$ antenna (*right*), with the sector resolution according to eq. (3.27). The sector accounting for a as the emitted gluon, shown in the right pane, is not considered during the shower evolution, as emission into the initial state is not accounted for.

3.3 Choice of sector resolution variables

As the antenna phase space is divided into sectors corresponding to the radiation from different antennae, cf. figures 3 and 4, a criterion to decide which branching $IK \mapsto ijk$ to perform has to be chosen. When thinking in terms of inverting the shower, this variable then determines which clustering $ijk \mapsto IK$ to perform. The obvious choice would be to simply chose the ordering variables eqs. (2.63) to (2.66), as its proportionality to $y_{ij}y_{jk}$ ensures that the most singular sector is picked if in gluon emissions either the soft or a collinear singularity is approached.

However, this choice is not unique and for LL accuracy it must only ensure that the antenna function with the correct divergent terms is picked. Different choices will then have different subleading logarithmic behaviour.

We follow the choice made in [38] and take the dimensionful evolution variables $p_\perp$ to discriminate between sectors for gluon emissions and a $p_\perp$-weighted virtuality for gluon splittings, as summarised in eqs. (3.25) to (3.28). We refrain from choosing the dimensionless equivalents, as it was shown in [38] to yield worse subleading-logarithmic behaviour for final-state gluon emissions.

JHEP07(2020)032

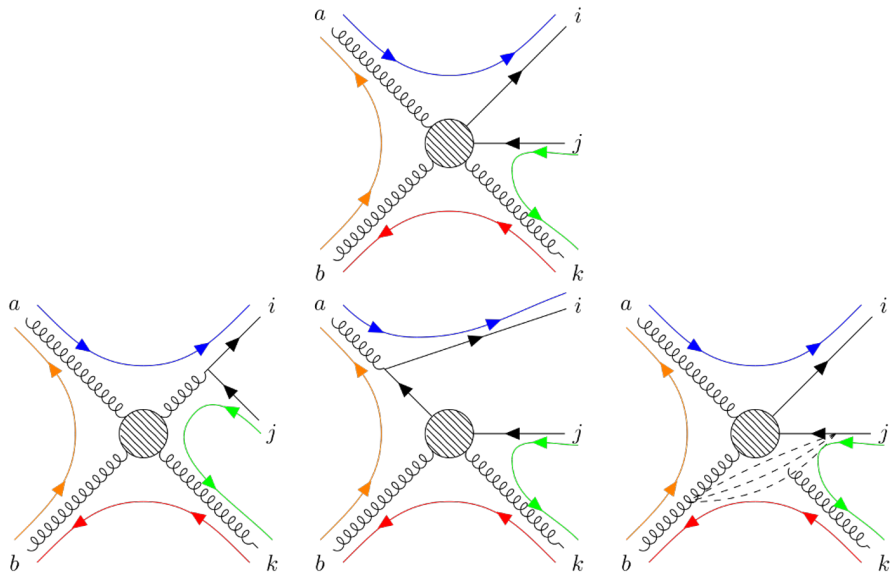


Figure 5. Illustration of all different histories (*bottom row*) leading to the same colour-ordered process $gg \to q\bar{q}g$ (*top row*). Going from left to right we have: a final-state gluon splitting, an initial state gluon splitting, and an initial-final gluon emission.

The particular choice for quark pairs can be understood by considering the colour-ordered process $gg \to q\bar{q}g$, with a colour-ordering as shown in figure 5. Three different histories contribute to this final state. Firstly, the emission of gluon k could have happened before or after the creation of the q-$\bar{q}$ pair. Secondly, this splitting could have both happened in the initial or final state. In total, this leads to four different possible clusterings,

- final-state gluon splitting: $ijk \mapsto IK$, $aij \mapsto AJ$
- initial-state gluon splitting: $aib \mapsto AB$
- gluon emission: $jkb \mapsto JB$

However, while the $\bar{q}_j$-g_k-g_b antenna contains the q_j-g_k-collinear singularity, the q_i-$\bar{q}_j$-g_k antenna only contains the quark-antiquark-collinear singularity and is finite for q_j and g_k becoming collinear. Hence, by comparing only the transverse momentum $p_\perp$ for $y_{jk} \to 0$, one is prone to pick the wrong clustering, if the invariant on the i-j-side is smaller than the one on the k-b-side. Using an interpolation between the geometric mean of the transverse momentum $p_\perp$ and the virtuality of the quark pair as collected below in eqs. (3.25) to (3.28) reflects this subtlety and guarantees that the entire $\bar{q}_j$-g_k-collinear limit is classified to belong to the $\bar{q}_j$-g_k-g_b antenna.

FF branchings. For final-final configurations, we choose the sector resolution variables:

$$Q^2_{\text{res}_j\text{FF}} = \begin{cases} \dfrac{s_{ij}s_{jk}}{s_{IK}} & \text{if } j \text{ is a gluon} \\ (s_{ij} + 2m_j^2)\sqrt{y_{jk} + \mu_j^2} & \text{if } (i,j) \text{ is a quark-antiquark pair} \end{cases}, \tag{3.25}$$

with the respective choice for $i \leftrightarrow k$.

JHEP07(2020)032

RF branchings. The RF sector resolution variables are virtually the same as the FF ones,

$$Q^2_{\mathrm{res}_j\mathrm{RF}} = \begin{cases} \dfrac{s_{aj}s_{jk}}{s_{aj}+s_{ak}} & \text{if } j \text{ is a gluon} \\ (s_{jk}+2m_j^2)\sqrt{y_{aj}-\mu_j^2} & \text{if } (j,k) \text{ is a quark-antiquark pair} \end{cases}, \tag{3.26}$$

with the difference only in the sign of the mass correction to y_{aj}, due to the different momentum conservation, cf. eq. (2.6).

IF branchings. For IF configurations, we choose the sector resolution variables to be the same as the RF ones, complemented by additional ones treating initial-state gluon splittings and conversions, not present in RF branchings:

$$Q^2_{\mathrm{res}_j\mathrm{IF}} = \begin{cases} \dfrac{s_{aj}s_{jk}}{s_{aj}+s_{ak}} & \text{if } j \text{ is a gluon} \\ (s_{aj}-2m_j^2)\sqrt{y_{jk}+\mu_j^2} & \text{if } (a,j) \text{ is a quark-quark pair} \\ s_{aj}\sqrt{y_{jk}+\mu_j^2} & \text{if } (a,j) \text{ is a gluon-(anti)quark pair} \\ (s_{jk}+2m_j^2)\sqrt{y_{aj}-\mu_j^2} & \text{if } (j,k) \text{ is a quark-antiquark pair} \end{cases}. \tag{3.27}$$

II branchings. For initial-initial configurations, we choose the resolution variables

$$Q^2_{\mathrm{res}_j\mathrm{II}} = \begin{cases} \dfrac{s_{aj}s_{jb}}{s_{ab}} & \text{if } j \text{ is a gluon} \\ (s_{aj}-2m_j^2)\sqrt{y_{jb}-\mu_j^2} & \text{if } (a,j) \text{ is a gluon-(anti)quark pair} \\ s_{aj}\sqrt{y_{jb}-\mu_j^2} & \text{if } (a,j) \text{ is a quark-quark pair} \end{cases}, \tag{3.28}$$

with the respective choice for $a \leftrightarrow b$.

4 Validation and preliminary results

We validate the sector shower in two stages: first in section 4.1 by comparing its tree-level expansion to tree-level matrix elements (see [12, 38, 49] for equivalent plots for global showers), and second in section 4.2 by comparing the full-fledged sector shower to experimental data. For the latter, comparisons are also included to the global Vincia and default Pythia 8.3 showers. Both Rivet [58] and an internal Vincia analysis package were used for performing analyses.

JHEP07(2020)032

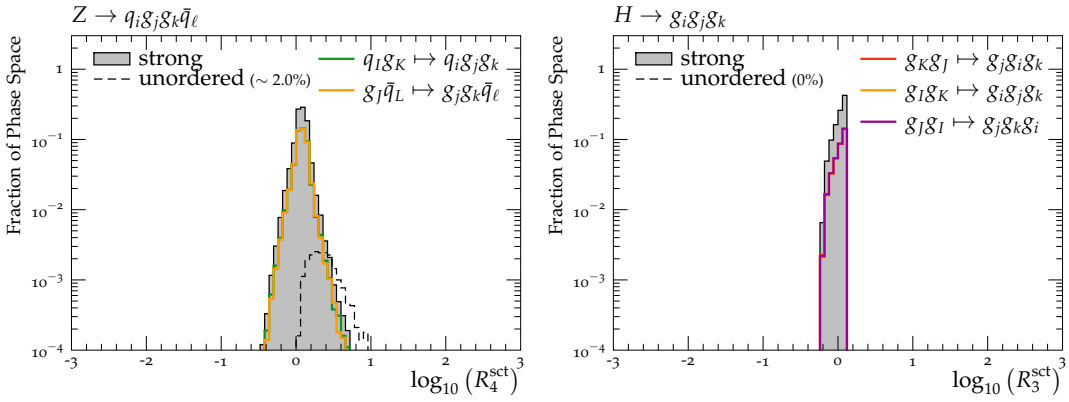


Figure 6. Ratios R_n^{sct} of the sector-shower approximation to LO matrix elements for $Z \to q\bar{q} + 2g$ and $H \to 3g$ in a flat phase space scan.

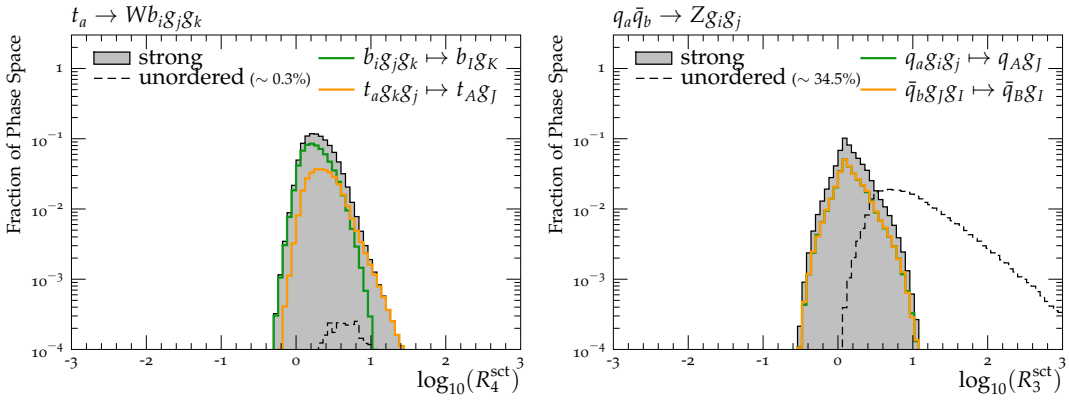


Figure 7. Ratios R_n^{sct} of the sector-shower approximation to LO matrix elements for gluon emissions off $t \to W^+ b$ (*left*) and $q\bar{q} \to Z$ (*right*) in a flat phase space scan.

4.1 Comparison to tree-level matrix elements

To probe the quality of our sector antenna functions as well as our sector resolution criteria, we define the parton-shower-to-matrix-element ratio

$$\begin{aligned} R_n^{\text{sct}} &= \left(\prod_{\substack{\text{ordered} \\ i}}^{n-n_{\text{Born}}} \sum_{j \in \{p\}_{n-i}} \Theta(Q^2_{\text{min}}(\{p\}_{n-i}) - Q^2_{\text{res}_j}) g_{\text{s}}^2 \mathcal{C}_j \bar{a}_j^{\text{sct}} \right) \frac{|\mathcal{M}_{\text{Born}}|^2}{|\mathcal{M}_n|^2} \\ &= \left(\prod_{\substack{\text{ordered} \\ i}} g_{\text{s}}^2 \mathcal{C}_i \bar{a}_i^{\text{sct}} \right) \frac{|\mathcal{M}_{\text{Born}}|^2}{|\mathcal{M}_n|^2} \,, \end{aligned} \tag{4.1}$$

where $Q^2_{\text{min}}(\{p\}_{n-i})$ is the minimal sector resolution over the state $\{p\}_{n-i}$ and the symbol $\prod_{\text{ordered}}$ is meant to impose strong ordering, i.e., that $p^2_{\perp,n} < p^2_{\perp,n-1} < p^2_{\perp,n-2} < \ldots < Q^2_{\text{F}}$ for each clustering sequence, where Q^2_{F} denotes the factorisation scale. By construction, *only a single clustering* contributes at each n-parton phase space point, as indicated by the second line in eq. (4.1).

JHEP07(2020)032

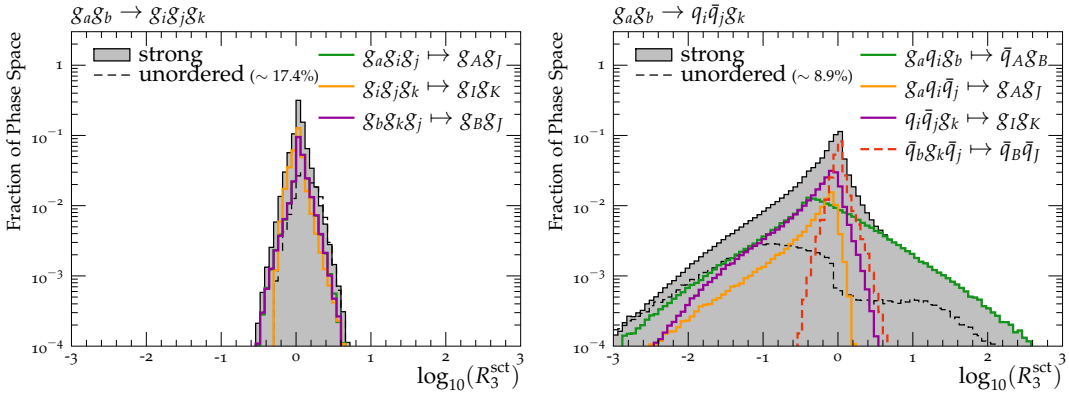


Figure 8. Ratios R_3^{sct} of the sector-shower approximation to LO matrix elements for gg-fusion processes in a flat phase space scan.

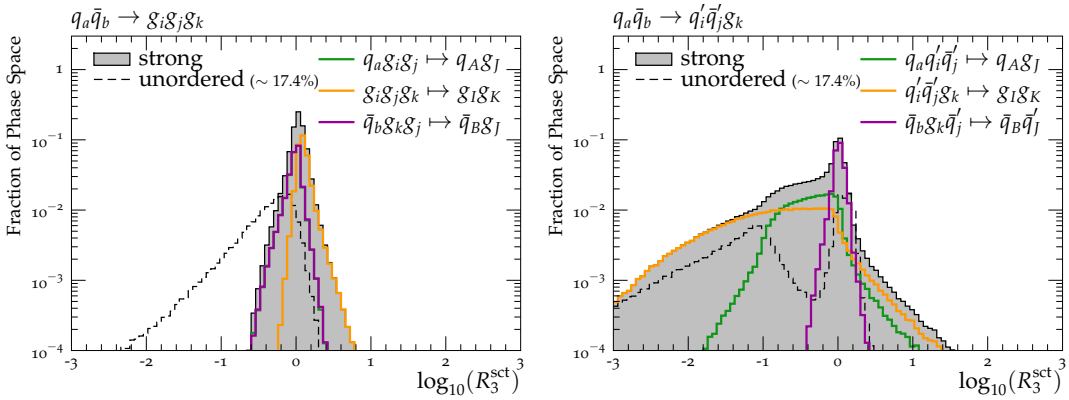


Figure 9. Ratios R_n^{sct} of the sector-shower approximation to LO matrix elements for $q\bar{q}$-annihilation processes in a flat phase-space scan.

For a fixed colour ordering, we compare the sector antenna shower approximation to leading-colour tree-level matrix elements from the MadGraph5_aMC@NLO matrix element generator [59]. Using Vincia's implementation of Rambo [60], we generate large samples of the n-parton phase space in a flat phase-space scan. Initial-state momenta are fixed by a flat sampling of the momentum fractions for partonic scattering processes and to the pole mass for resonance decays. We cluster the state back to the Born configuration according to the exact inverses of the corresponding $2 \mapsto 3$ kinematic maps, cf. section 2.3, and determine the sector-shower history according to the resolution variables eqs. (3.25) to (3.28).

In figure 6, we plot R_n^{sct} for the resonance decay processes $Z \to qgg\bar{q}$ and $H \to ggg$, while in figure 7, we show it for gluon emissions off the resonance decay $t \to W^+ b$ and the Drell-Yan process $q\bar{q} \to Z$. In figures 8 and 9, we consider $2 \to 3$ gluon-gluon and quark-antiquark scattering processes respectively. We here denote light (i.e. effectively massless) quarks by q and denote heavy flavours explicitly by t and b; unequal flavours are denoted by dashes, q'. For all cases, we show the shower approximations in both the ordered (solid) and unordered (dashed) regions of phase space. (The latter is shown for completeness; it would be vetoed in a strongly ordered shower.) In the ordered regions, we also include the

JHEP07(2020)032

breakdown per contributing sector (coloured lines). In figures 15 to 20 in appendix D, we present R_n^{sct} for higher multiplicities.

The patterns we observe are the following. For gluon emissions, we see quite narrow distributions centred on $\log(R_n) = 0$, indicating a generally very good agreement between the sector-shower approximation and the LO matrix elements. This is particularly true in the ordered regions of the phase space for each hard process. Branching processes involving gluon splittings (including initial-state splittings and quark conversions) are significantly less well approximated. This is due to a combination of two factors, as follows. Firstly, gluon-splitting kernels only contain single poles which hence are expected to dominate over a smaller fraction of phase space compared with the double poles of gluon-emission ones. Secondly, the numerator of the single poles of the $g \to q\bar{q}$ branchings has a significant dependence on the (linear) polarisation of the splitting gluon (which is currently not accounted for in our sector shower), while the polarisation dependence of gluon-emission processes is much milder [61]. The overall quality of the approximations are consistent with those seen for global showers in [12, 49] and for the final-final sector shower in [38].

Furthermore, for several of the QCD scattering processes, we observe that even the distributions of *unordered* histories are peaked close to $R_n^{\text{sct}} = 1$. Many of these "good" unordered paths are, however, only unordered in the first emission, i.e., the first emission appears at a higher scale than the factorisation scale. Although this suggests to let the shower evolution start at a nominally larger scale than the factorisation scale, we refrain from doing so for now and leave a study of such "power-shower" effects, cf. [62], in sector showers to future work.

4.2 Comparison to experimental data

We here present first results with the VINCIA sector shower, comparing it to experimental measurements and to the global VINCIA and PYTHIA 8.3 showers. Although we included a large set of observables in our study, we here show only a minimal set which were deemed to be the most physically relevant and/or representative ones.

With the splitting kernels and kinematics maps fixed, the main quantity governing the perturbative shower evolution is the strong coupling. For both of the VINCIA showers, we evolve the coupling at two-loop order with an $\overline{\text{MS}}$ value of $\alpha_{\text{s}}(m_Z) = 0.118$. Since our shower model is coherent, we take the CMW scheme [63] as the baseline scheme for defining the effective value of α_s for shower branchings, i.e., we use

$$\alpha_{\text{s}}^{\text{CMW}} = \alpha_{\text{s}}^{\overline{\text{MS}}} \left(1 + \frac{\alpha_{\text{s}}^{\overline{\text{MS}}}}{2\pi} \left[C_A \left(\frac{67}{18} - \frac{\pi^2}{6} \right) - \frac{5 n_f}{9} \right] \right) , \tag{4.2}$$

which is technically imposed via an n_f-dependent rescaling of the Λ_{QCD} parameter. The CMW factor is derived for gluon emissions in the infinitely soft limit. To reflect remaining ambiguities in the scale definition and to allow the effective value of α_s used for physical branchings to deviate from this, we further allow renormalisation-scale prefactors k_R to modify the evolution-scale $p_\perp$ argument of the running coupling, so that:

$$\alpha_s^{\text{Vincia}}(p_\perp^2) = \alpha_s^{\text{CMW}}(k_R p_\perp^2) \, . \tag{4.3}$$

JHEP07(2020)032

From preliminary tuning studies comparing VINCIA's global shower to LEP event shapes and Drell-Yan $p_\perp$ spectra, the default values for the additional scale prefactors are chosen as

$$k^{\mathrm{F}}_{\mathrm{R,Emit}} = 0.66\,, \quad k^{\mathrm{F}}_{\mathrm{R,Split}} = 0.8\,, \tag{4.4}$$

$$k^{\mathrm{I}}_{\mathrm{R,Emit}} = 0.66\,, \quad k^{\mathrm{I}}_{\mathrm{R,Split}} = 0.5\,, \quad k^{\mathrm{I}}_{\mathrm{R,Conv}} = 0.5\,. \tag{4.5}$$

We note that these choices roughly double the effect of the CMW scheme translation. We plan to return to an investigation of the seemingly large $1/k_R$ values needed to reproduce data in a future study including higher-order virtual effects (e.g., along the lines of [41]) but for the moment note that the combined effect is similar to that of the rather large effective value of $\alpha_s(M_Z) = 0.1365$ chosen in the baseline tune of PYTHIA [64].

The complete set of parameters changed relative to the default PYTHIA tune are collected in appendix E.

Electron-positron annihilation. As a first test of the sector shower implementation and specifically as a test of its FSR component, in figure 10 we show event-shape distributions for hadronic Z decays (see [67–72] for definitions) compared to L3 measurements at $\sqrt{s} = 91.2\,\mathrm{GeV}$ [65] for light- and heavy-flavour tagged samples separately.

Although we are considering only pure shower predictions here, i.e., without matching or merging to fixed-order calculations, we find that both the sector and global showers in VINCIA give very good agreement with data. All three showers considered here describe the thrust and C parameter distributions equally well; the sector-shower predictions in decays to light flavours are in slightly closer agreement with data than the global shower. The default PYTHIA 8.3 shower is slightly closer to the experimentally measured distributions for the D parameter in Z decays to light flavours. We note, however, that its first emission is corrected to LO matrix elements and the second includes an approximate treatment of gluon-polarisation effects, whereas no such corrections are imposed on the VINCIA showers. Given that the latter two agree with each other, we suspect this difference to mainly stem from the additional corrections that are implemented in PYTHIA. This argument is supported by a study of the distribution obtained with the unmatched PYTHIA 8.3 shower and the fact that the LO matched VINCIA 1 shower agreed better with data as well as PYTHIA, cf. [12]. We plan to return to (iterative) matrix-element corrections and (multi-leg) matching/merging for VINCIA in a separate follow-up study.

Drell-Yan. To test the ISR component of the sector shower with minimal interplay with the FSR shower, we consider Drell-Yan processes as measured by the D∅ experiment [66] at $\sqrt{s} = 1.96$ TeV in figure 11. We study the angle ϕ^*_η, defined in [74], which relates to the opening angle of the Drell-Yan lepton pair and is sensitive to the Z transverse momentum.

It is known that a pure unmatched parton shower with the evolution starting at the factorisation scale $\sim m_Z$ is incapable of describing these observables well, as they are sensitive to phase-space regions with $p_{\perp,\mathrm{jet}} > m_Z$. We here follow the default choice in both PYTHIA and VINCIA and start showers off configurations containing no final-state (QCD) partons at the phase-space maximum. As alluded to before, cf. section 4.1, we leave a dedicated study of the impact of so-called "power showers" in sector showers as well as matching and merging in the sector-shower framework to future work.

JHEP07(2020)032

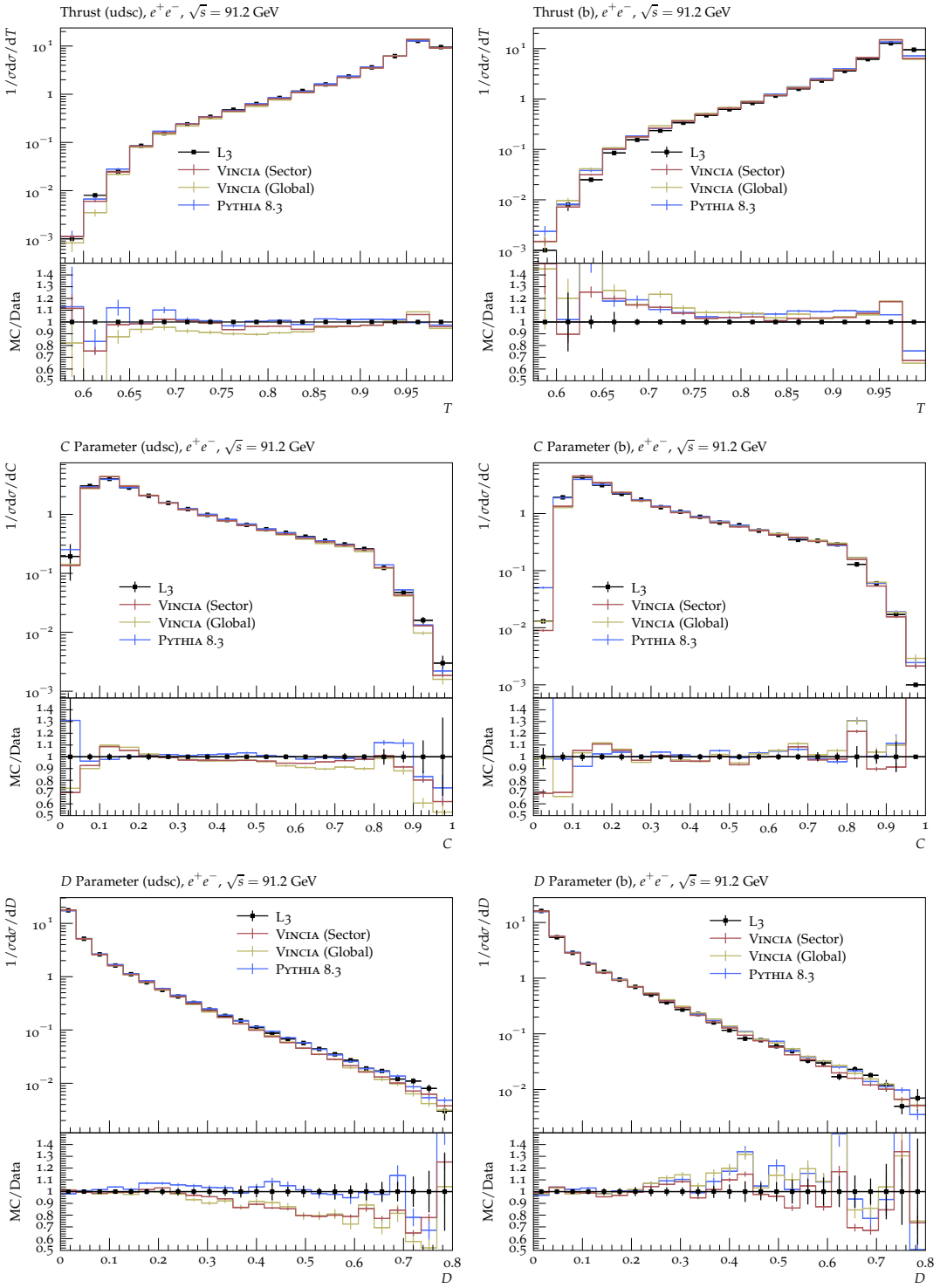


Figure 10. Normalised event shape distributions in Z decays to light flavours (*left*) and b quarks only (*right*) in e^+e^- collisions at a centre-of-mass energy of $\sqrt{s} = 91.2$ GeV. Comparison of the sector shower against the default (global) VINCIA shower, PYTHIA 8.3, and L3 data [65].

JHEP07(2020)032

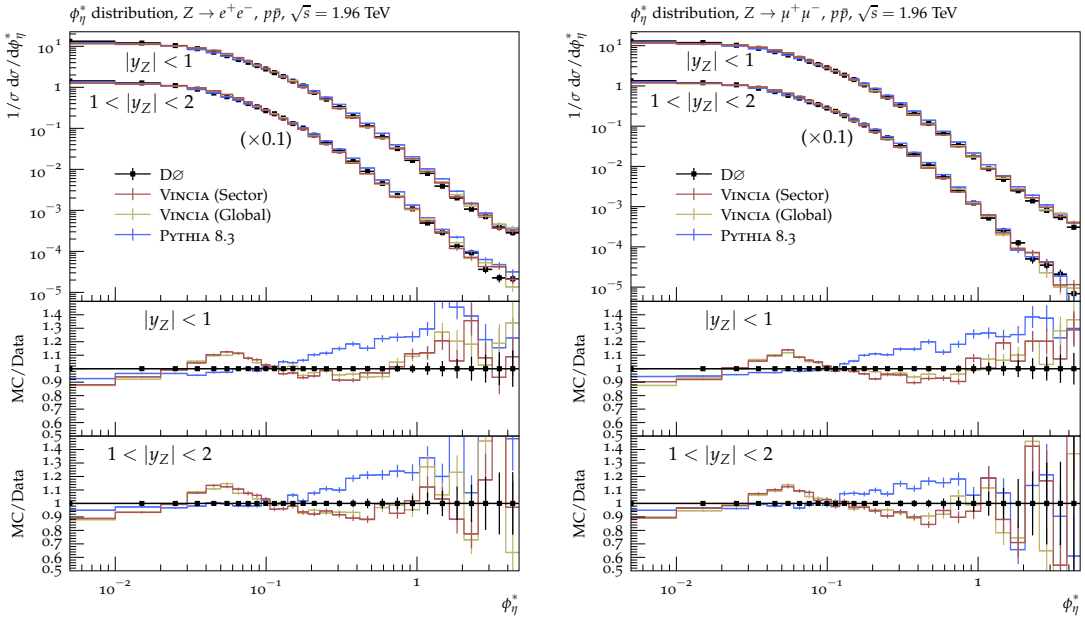


Figure 11. Distribution of the ϕ^*_η angle in the electron channel (*left*) and muon channel (*right*) for different rapidity bins in $p\bar{p}$ collisions at a centre-of-mass energy of $\sqrt{s} = 1.96$ TeV. Comparison of the sector shower against the default (global) Vincia shower, Pythia 8.3 and DØ data [66].

With this choice, the data is very well described by the sector shower, although there is some larger deviation for large ϕ^*. This is, however, the region a "wimpy" shower starting at the factorisation scale would not describe well at all and fixed-order corrections have a great impact, cf. e.g. [19, 49]. A peculiarity of Drell-Yan processes is that the first emission is described in the same way in the sector and global shower framework in Vincia, as initial-state legs are sectorised in the first place. Only starting from the second branching do different initial-final sectors compete, and final-final ones do not show up until the third branching. This is observed in figure 11, as there is virtually no difference between the predictions of the sector and global shower.

QCD jets. As a test of the complete sector shower with full interplay between its ISR and FSR components, we compare sector-shower predictions for dijet azimuthal decorrelations in proton-proton collisions at $\sqrt{s} = 7$ TeV to measurements by the ATLAS experiment [73] in figure 12.

Of all analyses considered here, the different shower predictions differ the most for this observable. Systematically, the sector shower is shifted towards smaller values compared to the global shower. However, already for the first emission, many different sectors compete for the branching in the sector shower, cf. section 4.1, which may explain this bigger difference to the global shower than seen in the pure FSR shower and showers off Drell-Yan processes. Moreover, dijet azimuthal decorrelations are sensitive to higher-order contributions and typically, NLO theory predictions are required to describe the data well, cf. [76]; here however only shower predictions matched to at most LO for the first emission are considered. The good agreement that we nevertheless see may be explained by noting that the distributions are normalised to the cross section.

JHEP07(2020)032

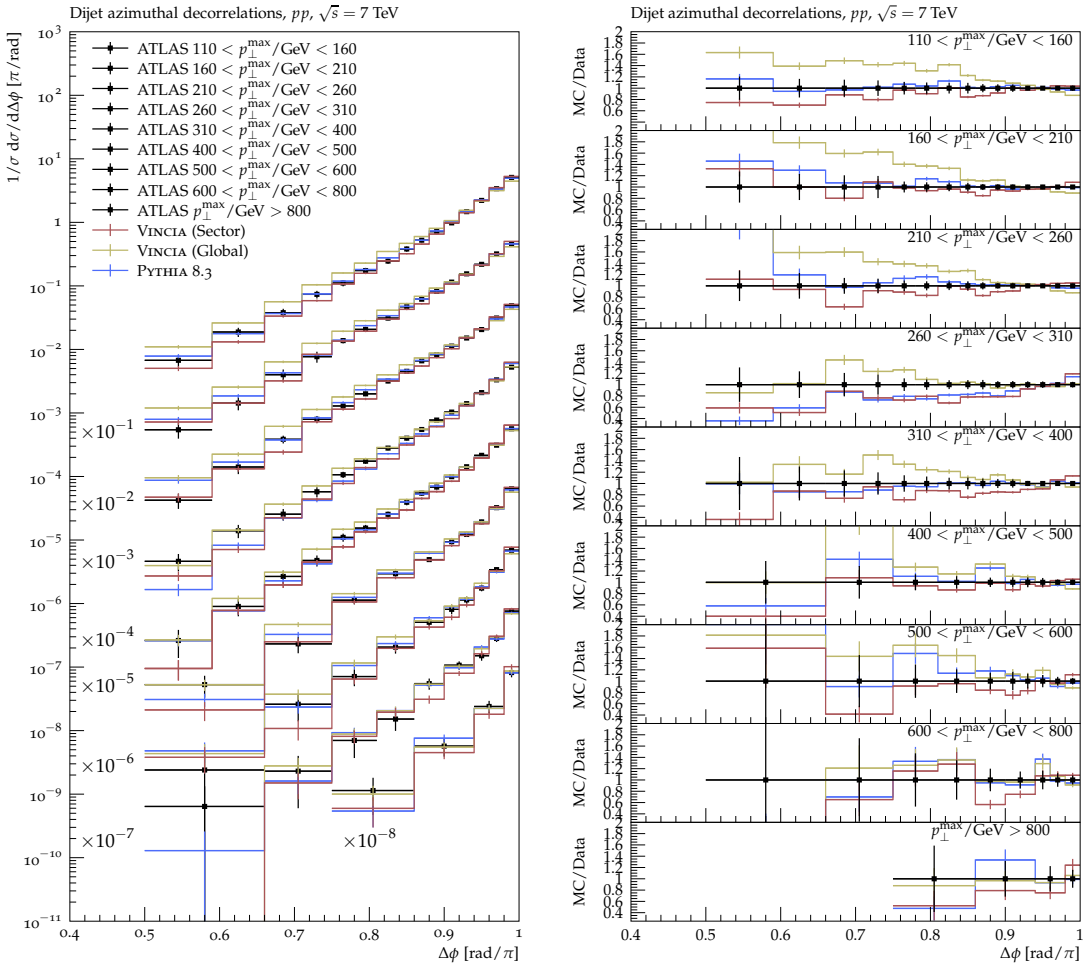


Figure 12. Dijet azimuthal decorrelations in pp collisions at a centre-of-mass energy of $\sqrt{s} = 7$ TeV. Comparison of the sector shower against the default (global) Vincia shower, Pythia 8.3, and ATLAS data [73].

Coloured resonance decays. As a final validation of the sector shower in resonance decays, we consider two observables already studied in [50], known to be sensitive to RF branchings. The first observable is the b-jet profile $\rho(r)$ as measured by ATLAS for $pp \to t\bar{t}$ at $\sqrt{s} = 7\,\mathrm{TeV}$ [75]. As shown in figure 13 we find that the sector shower reproduces the global shower results, namely that Vincia gives rise to systematically narrower b-jets than Pythia, and moreover that this is in closer agreement with the data.

The second observable we consider is the invariant mass of the $\bar{t}$ decay system $\bar{b}_j \ell^- \bar{\nu}_\ell$ in dileptonic $pp \to t\bar{t}$ at $\sqrt{s} = 8\,\mathrm{TeV}$. As already noted in [50], this observable is sensitive to numerous effects; therefore to isolate (and thereby highlight) any differences originating from the parton shower alone, we do the following. The analysis is performed at "Monte Carlo truth" level, namely the decay system is reconstructed with the correct pairing of leptons and b-jets, and the four-momentum of the neutrino is presumed to be known. In ad-

JHEP07(2020)032

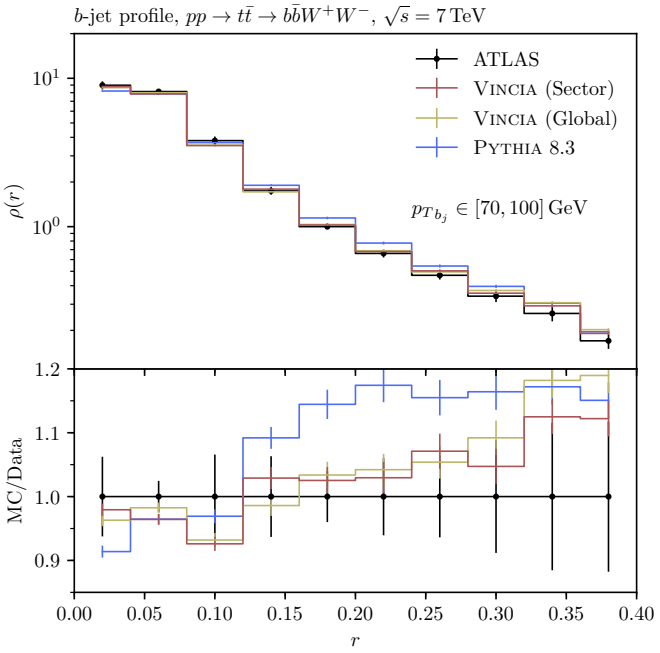


Figure 13. Comparison of the sector shower against the default (global) VINCIA shower, and PYTHIA 8.3 for the b-jet profile as measured by ATLAS in $pp \to t\bar{t}$ at $\sqrt{s} = 7\,\text{TeV}$ [75].

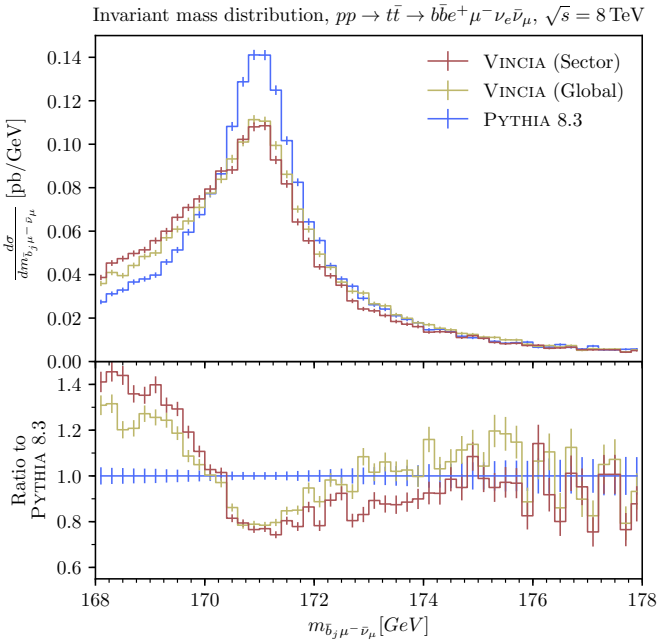


Figure 14. Comparison of the sector shower against the default (global) VINCIA shower, and PYTHIA 8.3 for invariant mass distribution of the $\bar{t}$ decay system $\bar{b}_j\mu^-\bar{\nu}_\mu$ in dileptonic $pp \to t\bar{t}$ at $\sqrt{s} = 8\,\text{TeV}$.

JHEP07(2020)032

dition, we switch off the simulation of MPI, beam-remnant modelling, and hadronisation.[12] These actions strengthen our conclusion that the sector shower is consistent with the global shower, as demonstrated in figure 14. Although there is again a small systematic shift of the sector shower towards smaller invariant masses relative to the global shower, this is well within an acceptable range of variation for showers of leading-logarithmic accuracy. Indeed, the difference between the two VINCIA showers is significantly smaller than their mutual difference with respect to the predictions of PYTHIA 8.3. The variation between VINCIA and PYTHIA is interesting. It was discussed in some detail in [50]; here we produce a brief summary. In the PYTHIA shower model, the recoil from each FSR emission must be taken by a single final-state parton. For the first gluon emission in $t \to bW$, therefore, the W is formally assigned to be the recoiler for emissions off the b quark. At this stage, this choice is fully equivalent to VINCIA's RF kinematics (since $p_t - p_b = p_W$). After the first branching, the colour index of the top quark is now carried by the emitted gluon. In VINCIA, the "recoiler" for further emissions off that colour index effectively remains the crossed top minus the radiating parton, $p_t - p_g = p_W + p_b$. In PYTHIA, instead, one must choose it to be either the W or the b. This choice is controlled by the flag `RecoilToColoured` whose default value `true` means that in ambiguous cases like this one the coloured parton will be picked over the uncoloured one. (This is to avoid imparting undesirably large recoil effects from QCD branchings onto non-QCD partons.) Due to the collinear enhancement for the preceding $b \to bg$ emission, however, the invariant mass of the bg dipole (and hence the phase space for further radiation) tends to be small. This severely restricts further radiation and results in less "out-of-cone" radiation loss in PYTHIA and hence a sharper peak structure in figure 14. It should be clear from our discussion that we consider this effect in PYTHIA to be an artefact of having to choose between two undesirable options; either allowing an uncoloured parton to continue to receive the full recoil of further QCD branchings in the RF dipole (`RecoilToColoured = false`) or choosing a coloured one with a tiny phase space (`RecoilToColoured = true`) to receive it. We believe that VINCIA's coherent treatment of RF recoils is more physical. We also note that, since this difference arises in the treatment of the second emission and beyond, it is of NNLO origin in a fixed-order expansion and will persist even when the decay process is matched to NLO calculations, see [50, 77].

5 Conclusion and outlook

We here presented the first implementation of an ISR and FSR parton shower that possesses only a single shower history. We extensively reviewed the shower algorithm which is ordered in a generalised ARIADNE transverse momentum and presented the construction of all required sector antenna functions from VINCIA's global final-final antenna functions; we also defined a sector resolution variable, which enables us to uniquely invert the shower algorithm.

The sector-shower implementation was validated against LO matrix elements and first predictions for physical observables were made and compared to the global VINCIA antenna

[12] In addition, we set `SigmaProcess:alphaSorder` and `SigmaProcess:alphaSvalue` to their default values in PYTHIA (such that the starting cross section is identical).

JHEP07(2020)032

shower, and to PYTHIA 8.3. Where relevant, we emphasised the expected impact of matrix-element corrections, which are available in PYTHIA e.g. for resonance decays and Drell-Yan processes, but are not yet available in the current version of VINCIA.

We found very good agreement of sector-shower predictions with experimental data from LEP, Tevatron, and LHC experiments, taking into account that only "pure shower" predictions, formally at LL accuracy, were considered. Moreover, we showed that, at least for the observables studied here, there is no disadvantage of the sector shower in describing data in comparison to the global VINCIA shower.

While this may appear to be "just another LL shower", it is worth pointing out that the sector-shower approach provides the means to significantly reduce the computational overhead connected to matching and merging shower predictions to fixed-order calculations and, moreover, to significantly reduce the complexity of shower algorithms. The presentation here may therefore be understood as a first step towards dedicated and highly-efficient matrix-element-correction and merging schemes, which we discuss in separate forthcoming publications. We also plan to explore whether the sector approach can facilitate the inclusion of direct higher-order splittings and one-loop antenna functions in the shower, e.g. following a similar approach as those of [41, 42], with the ultimate goal of reaching higher formal precision in parton showers, cf. [78–83].

The sector antenna shower presented here will be made publicly available as part of the VINCIA antenna shower in a forthcoming update to PYTHIA 8.3.

Acknowledgments

We acknowledge support from the Monash eResearch Centre and eSolutions-Research Support Services through the MonARCH HPC Cluster. CTP is supported by the Monash Graduate Scholarship, the Monash International Postgraduate Research Scholarship, and the J. L. William Scholarship. HB is funded by the Australian Research Council via Discovery Project DP170100708 — "Emergent Phenomena in Quantum Chromodynamics". This work was also supported in part by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 722104 — MCnetITN3.

A Helicity-dependent antenna functions

In this section, we collect all helicity- and mass-dependent global antenna functions. The corresponding sector antenna functions can be build by the considerations presented in section section 3.1.

A.1 Final-final antennae

QQEmitFF. The helicity-averaged antenna function for the process $q_I q_K \mapsto q_i g_j q_k$ is

$$\bar{a}^{\mathrm{FF,gl}}_{g/q\bar{q}}(y_{ij}, y_{jk}) = \frac{1}{s_{IK}} \left[\frac{(1-y_{ij})^2 + (1-y_{jk})^2}{y_{ij} y_{jk}} - \frac{2\mu_I^2}{y_{ij}^2} - \frac{2\mu_K^2}{y_{jk}^2} + 1 \right] \tag{A.1}$$

JHEP07(2020)032

with the individual helicity contributions

$$\bar{a}^{\mathrm{FF,gl}}_{q_+\bar{q}_+\mapsto q_+g_+\bar{q}_+} = \frac{1}{s_{IK}}\left[\frac{1}{y_{ij}y_{jk}} - \frac{\mu_i^2}{y_{ij}^2(1-y_{jk})} - \frac{\mu_k^2}{y_{jk}^2(1-y_{ij})}\right], \tag{A.2}$$

$$\bar{a}^{\mathrm{FF,gl}}_{q_+\bar{q}_+\mapsto q_+g_-\bar{q}_+} = \frac{1}{s_{IK}}\left[\frac{(1-y_{ij})^2+(1-y_{jk})^2-1}{y_{ij}y_{jk}} + 2 - \frac{\mu_i^2(1-y_{jk})}{y_{ij}^2} - \frac{\mu_k^2(1-y_{ij})}{y_{jk}^2}\right], \tag{A.3}$$

$$\bar{a}^{\mathrm{FF,gl}}_{q_+\bar{q}_+\mapsto q_-g_+\bar{q}_+} = \frac{1}{s_{IK}}\frac{\mu_i^2 y_{jk}^2}{y_{ij}^2}\frac{1}{1-y_{jk}}, \tag{A.4}$$

$$\bar{a}^{\mathrm{FF,gl}}_{q_+\bar{q}_+\mapsto q_+g_+\bar{q}_-} = \frac{1}{s_{IK}}\frac{\mu_k^2 y_{ij}^2}{y_{jk}^2}\frac{1}{1-y_{ij}}, \tag{A.5}$$

$$\bar{a}^{\mathrm{FF,gl}}_{q_+\bar{q}_-\mapsto q_+g_+\bar{q}_-} = \frac{1}{s_{IK}}\left[\frac{(1-y_{ij})^2}{y_{ij}y_{jk}} - \frac{\mu_i^2}{y_{ij}^2(1-y_{jk})} - \frac{\mu_k^2(1-y_{ij})}{y_{jk}^2}\right], \tag{A.6}$$

$$\bar{a}^{\mathrm{FF,gl}}_{q_+\bar{q}_-\mapsto q_+g_-\bar{q}_-} = \frac{1}{s_{IK}}\left[\frac{(1-y_{jk})^2}{y_{ij}y_{jk}} - \frac{\mu_i^2(1-y_{jk})}{y_{ij}^2} - \frac{\mu_k^2}{y_{jk}^2(1-y_{ij})}\right], \tag{A.7}$$

$$\bar{a}^{\mathrm{FF,gl}}_{q_+\bar{q}_-\mapsto q_-g_+\bar{q}_-} = \frac{1}{s_{IK}}\frac{\mu_i^2 y_{jk}^2}{y_{ij}^2(1-y_{jk})}, \tag{A.8}$$

$$\bar{a}^{\mathrm{FF,gl}}_{q_+\bar{q}_-\mapsto q_+g_-\bar{q}_+} = \frac{1}{s_{IK}}\frac{\mu_k^2 y_{ij}^2}{y_{jk}^2(1-y_{ij})}. \tag{A.9}$$

QGEmitFF. The helicity-averaged antenna function for the process $q_I g_K \mapsto q_i g_j g_k$ is

$$\bar{a}^{\mathrm{FF,gl}}_{g/qg}(y_{ij},y_{jk}) = \frac{1}{s_{IK}}\left[\frac{(1-y_{ij})^3+(1-y_{jk})^2}{y_{ij}y_{jk}} - \frac{2\mu_I^2}{y_{ij}^2} + (1-\alpha)\frac{(1-y_{jk})(y_{ik}-y_{ij})}{y_{jk}} + 2 - y_{ij} - \frac{y_{jk}}{2}\right] \tag{A.10}$$

with the individual helicity contributions

$$\bar{a}^{\mathrm{FF,gl}}_{q_+g_+\mapsto q_+g_+g_+} = \frac{1}{s_{IK}}\left[\frac{1}{y_{ij}y_{jk}} + (1-\alpha)(1-y_{jk})\left(\frac{1-2y_{ij}-y_{jk}}{y_{jk}}\right) - \frac{\mu_i^2}{y_{ij}^2(1-y_{jk})}\right], \tag{A.11}$$

$$\bar{a}^{\mathrm{FF,gl}}_{q_+g_+\mapsto q_+g_-g_+} = \frac{1}{s_{IK}}\left[\frac{(1-y_{ij})y_{ik}^2}{y_{ij}y_{jk}} - \frac{\mu_i^2(1-y_{jk})}{y_{ij}^2}\right], \tag{A.12}$$

$$\bar{a}^{\mathrm{FF,gl}}_{q_+g_+\mapsto q_-g_+g_+} = \frac{1}{s_{IK}}\frac{\mu_i^2 y_{jk}^2}{y_{ij}^2(1-y_{jk})}, \tag{A.13}$$

$$\bar{a}^{\mathrm{FF,gl}}_{q_+g_-\mapsto q_+g_+g_-} = \frac{1}{s_{IK}}\left[\frac{(1-y_{ij})^3}{y_{ij}y_{jk}} - \frac{\mu_i^2}{y_{ij}^2(1-y_{jk})}\right], \tag{A.14}$$

JHEP07(2020)032

$$\bar{a}^{\mathrm{FF,gl}}_{q_+g_-\mapsto q_+g_-g_-} = \frac{1}{s_{IK}}\left[\frac{(1-y_{jk})^2}{y_{ij}y_{jk}} + (1-\alpha)(1-y_{jk})\left(\frac{1-2y_{ij}-y_{jk}}{y_{jk}}\right) - \frac{\mu_i^2(1-y_{jk})}{y_{ij}^2}\right], \tag{A.15}$$

$$\bar{a}^{\mathrm{FF,gl}}_{q_+g_-\mapsto q_-g_+g_-} = \frac{1}{s_{IK}}\frac{\mu_i^2 y_{jk}^2}{y_{ij}^2(1-y_{jk})}. \tag{A.16}$$

GGEmitFF. The helicity-averaged antenna function for the process $g_I g_K \mapsto g_i g_j g_k$ is

$$\bar{a}^{\mathrm{FF,gl}}_{g/gg}(y_{ij},y_{jk}) = \frac{1}{s_{IK}}\left[\frac{(1-y_{ij})^3+(1-y_{jk})^3}{y_{ij}y_{jk}} + (1-\alpha)\frac{(1-y_{ij})(2y_{ik}-2y_{jk}+y_{ij})}{2y_{ij}} + (1-\alpha)\frac{(1-y_{jk})(2y_{ik}-2y_{ij}+y_{jk})}{2y_{jk}} + 3 - \frac{3}{2}y_{ij} - \frac{3}{2}y_{jk}\right] \tag{A.17}$$

with the individual helicity contributions

$$\bar{a}^{\mathrm{FF,gl}}_{g_+g_+\mapsto g_+g_+g_+} = \frac{1}{s_{IK}}\left[\frac{1}{y_{ij}y_{jk}} + (1-\alpha)\left((1-y_{ij})\frac{1-2y_{jk}-y_{ij}}{y_{ij}} + (1-y_{jk})\frac{1-2y_{ij}-y_{jk}}{y_{jk}}\right)\right] \tag{A.18}$$

$$\bar{a}^{\mathrm{FF,gl}}_{g_+g_+\mapsto g_+g_-g_+} = \frac{1}{s_{IK}}\frac{y_{ik}^3}{y_{ij}y_{jk}} \tag{A.19}$$

$$\bar{a}^{\mathrm{FF,gl}}_{g_+g_-\mapsto g_+g_+g_-} = \frac{1}{s_{IK}}\left[\frac{(1-y_{ij})^3}{y_{ij}y_{jk}} + (1-\alpha)(1-y_{ij})\frac{1-2y_{jk}}{y_{ij}}\right], \tag{A.20}$$

$$\bar{a}^{\mathrm{FF,gl}}_{g_+g_-\mapsto g_+g_-g_-} = \frac{1}{s_{IK}}\left[\frac{(1-y_{jk})^3}{y_{ij}y_{jk}} + (1-\alpha)(1-y_{jk})\frac{1-2y_{ij}}{y_{jk}}\right]. \tag{A.21}$$

GXSplitFF. The helicity-averaged antenna function for the process $g_I X_K \mapsto q_i \bar{q}_j X_k$ with an arbitrary coloured spectator X is

$$\bar{a}^{\mathrm{FF,gl}}_{\bar{q}/gX}(y_{ij},y_{jk}) = \frac{1}{s_{IK}}\frac{1}{2}\frac{1}{y_{ij}+2\mu_q^2}\left[y_{ik}^2 + y_{jk}^2 + \frac{2\mu_q^2}{y_{ij}+2\mu_q^2}\right] \tag{A.22}$$

with the individual helicity contributions

$$\bar{a}^{\mathrm{FF,gl}}_{g_+X\mapsto q_+\bar{q}_-X} = \frac{1}{s_{IK}}\frac{1}{2(y_{ij}+2\mu_q^2)}\left[y_{ik}^2 - \frac{\mu_q^2}{y_{ij}+2\mu_q^2}\frac{y_{ik}}{1-y_{ik}}\right], \tag{A.23}$$

$$\bar{a}^{\mathrm{FF,gl}}_{g_+X\mapsto q_-\bar{q}_+X} = \frac{1}{s_{IK}}\frac{1}{2(y_{ij}+2\mu_q^2)}\left[y_{jk}^2 - \frac{\mu_q^2}{y_{ij}+2\mu_q^2}\frac{y_{jk}}{1-y_{jk}}\right], \tag{A.24}$$

$$\bar{a}^{\mathrm{FF,gl}}_{g_+X\mapsto q_+\bar{q}_+X} = \frac{1}{s_{IK}}\frac{\mu_q^2}{2(y_{ij}+2\mu_q^2)^2}\left[\frac{y_{ik}}{1-y_{ik}} + \frac{y_{jk}}{1-y_{jk}} + 2\right]. \tag{A.25}$$

JHEP07(2020)032

A.2 Initial-final antennae

QQEmitIF. The helicity-averaged antenna function for the process $q_A q_K \mapsto q_a g_j q_k$ is

$$\begin{aligned}\bar{a}^{\mathrm{IF,gl}}_{g/qq}(y_{aj},y_{jk}) = \frac{1}{s_{AK}}\Bigg[&\frac{(1-y_{aj})^2+(1-y_{jk})^2}{y_{aj}y_{jk}} + \frac{1}{2}(2-y_{aj})(2-y_{jk}) \\ &-\frac{2\mu_a^2}{y_{aj}^2}\left((1-y_{jk})\left(1-\frac{y_{aj}}{2}\right)-\frac{y_{aj}}{2}(1-y_{aj})\right) \\ &-\frac{2\mu_k^2}{y_{jk}^2}\left(1-\frac{y_{jk}}{4}(2-y_{jk})\left(2+\frac{y_{aj}^2}{1-y_{aj}}\right)\right)\Bigg]\end{aligned} \tag{A.26}$$

with the individual helicity contributions

$$\bar{a}^{\mathrm{IF,gl}}_{q_+q_+\mapsto q_+g_+q_+} = \frac{1}{s_{AK}}\left[\frac{1}{y_{aj}y_{jk}} - \frac{\mu_a^2}{y_{aj}^2} - \frac{\mu_k^2}{(1-y_{aj})y_{jk}^2}\right], \tag{A.27}$$

$$\bar{a}^{\mathrm{IF,gl}}_{q_+q_+\mapsto q_+g_-q_+} = \frac{1}{s_{AK}}\left[\frac{(1-y_{aj})^2+[(1-y_{jk})^2-1](1-y_{aj})^2}{y_{aj}y_{jk}}\right. \tag{A.28}$$

$$\left. -\frac{\mu_a^2(1-y_{jk}-y_{aj})^2}{y_{aj}^2} - \frac{\mu_k^2(1-y_{aj})(1-y_{jk})^2}{y_{jk}^2}\right], \tag{A.29}$$

$$\bar{a}^{\mathrm{IF,gl}}_{q_+q_+\mapsto q_-g_-q_+} = \frac{1}{s_{AK}}\frac{\mu_a^2 y_{jk}^2}{y_{aj}^2}, \tag{A.30}$$

$$\bar{a}^{\mathrm{IF,gl}}_{q_+q_+\mapsto q_+g_+q_-} = \frac{1}{s_{AK}}\frac{\mu_k^2 y_{aj}^2}{(1-y_{aj})y_{jk}^2}, \tag{A.31}$$

$$\bar{a}^{\mathrm{IF,gl}}_{q_+q_-\mapsto q_+g_+q_-} = \frac{1}{s_{AK}}\left[\frac{(1-y_{aj})^2}{y_{aj}y_{jk}} - \frac{\mu_a^2(1-y_{aj})}{y_{aj}^2} - \frac{\mu_k^2(1-y_{aj})}{y_{jk}^2}\right], \tag{A.32}$$

$$\bar{a}^{\mathrm{IF,gl}}_{q_+q_-\mapsto q_+g_-q_-} = \frac{1}{s_{AK}}\left[\frac{(1-y_{jk})^2}{y_{aj}y_{jk}} - \frac{\mu_a^2(1-y_{jk})^2}{y_{aj}^2} - \frac{\mu_k^2(1-y_{jk})^2}{y_{jk}^2(1-y_{aj})}\right], \tag{A.33}$$

$$\bar{a}^{\mathrm{IF,gl}}_{q_+q_-\mapsto q_-g_-q_-} = \frac{1}{s_{AK}}\frac{\mu_a^2 y_{jk}^2}{y_{aj}^2}, \tag{A.34}$$

$$\bar{a}^{\mathrm{IF,gl}}_{q_+q_-\mapsto q_+g_-q_+} = \frac{1}{s_{AK}}\frac{\mu_k^2 y_{aj}^2}{y_{jk}^2(1-y_{aj})}. \tag{A.35}$$

QGEmitIF. The helicity-averaged antenna function for the process $q_A g_K \mapsto q_a g_j g_k$ is

$$\begin{aligned}\bar{a}^{\mathrm{IF,gl}}_{g/qg}(y_{aj},y_{jk}) = \frac{1}{s_{AK}}\Bigg[&\frac{(1-y_{aj})^3+(1-y_{jk})^2}{y_{aj}y_{jk}} + (1-\alpha)\frac{1-2y_{aj}}{y_{jk}} \\ &-\frac{2\mu_a^2}{y_{aj}^2}\left((1-y_{jk})-\frac{y_{aj}}{2}\left[(1-y_{aj})-(2-y_{jk})^2\right]\right) \\ &+\frac{3}{2}+y_{aj}-\frac{y_{jk}}{2}-\frac{y_{aj}^2}{2}\Bigg]\end{aligned} \tag{A.36}$$

JHEP07(2020)032

with the individual helicity contributions

$$\bar{a}^{\mathrm{IF,gl}}_{q_+g_+\mapsto q_+g_+g_+} = \frac{1}{s_{AK}}\left[\frac{1}{y_{aj}y_{jk}} + (1-\alpha)\frac{1-2y_{aj}}{y_{jk}} - \frac{\mu_a^2}{y_{aj}^2}\right], \tag{A.37}$$

$$\bar{a}^{\mathrm{IF,gl}}_{q_+g_+\mapsto q_+g_-g_+} = \frac{1}{s_{AK}}\left[\frac{(1-y_{aj})^3 + (1-y_{jk})^2 - 1}{y_{aj}y_{jk}}\right. \tag{A.38}$$

$$\left. - \frac{\mu_a^2(1-y_{jk}-y_{aj})^2(1-y_{aj})}{y_{aj}^2} + 3 - y_{aj}^2\right], \tag{A.39}$$

$$\bar{a}^{\mathrm{IF,gl}}_{q_+g_+\mapsto q_-g_-g_+} = \frac{1}{s_{AK}}\frac{\mu_a^2 y_{jk}^2}{y_{aj}^2}, \tag{A.40}$$

$$\bar{a}^{\mathrm{IF,gl}}_{q_+g_-\mapsto q_+g_+g_-} = \frac{1}{s_{AK}}\left[\frac{(1-y_{aj})^3}{y_{aj}y_{jk}} - \frac{\mu_a^2(1-y_{aj})^2}{y_{aj}^2}\right], \tag{A.41}$$

$$\bar{a}^{\mathrm{IF,gl}}_{q_+g_-\mapsto q_+g_-g_-} = \frac{1}{s_{AK}}\left[\frac{(1-y_{jk})^2}{y_{aj}y_{jk}} + (1-\alpha)\frac{1-2y_{aj}}{y_{jk}}\right. \tag{A.42}$$

$$\left. - \frac{\mu_a^2(1-y_{jk})^2}{y_{aj}^2} + 2y_{aj} - y_{jk}\right], \tag{A.43}$$

$$\bar{a}^{\mathrm{IF,gl}}_{q_+g_-\mapsto q_-g_-g_-} = \frac{1}{s_{AK}}\frac{\mu_a^2 y_{jk}^2}{y_{aj}^2}. \tag{A.44}$$

GQEmitIF. The helicity-averaged antenna function for the process $g_A q_K \mapsto g_a g_j q_k$ is

$$\bar{a}^{\mathrm{IF,gl}}_{g/gq}(y_{aj}, y_{jk}) = \frac{1}{s_{AK}}\left[\frac{(1-y_{jk})^3 + (1-y_{aj})^2}{y_{aj}y_{jk}} + \frac{1+y_{jk}^3}{y_{aj}(1-y_{jk})}\right. \tag{A.45}$$

$$- \frac{2\mu_k^2}{y_{jk}^2}\left(1 - \frac{y_{jk}}{4}(3 - 3y_{jk}^2 + y_{jk}^3)\left(2 + \frac{y_{aj}^2}{1-y_{aj}}\right)\right) \tag{A.46}$$

$$\left. + \frac{1}{2}(2-y_{aj})(3-y_{jk}+y_{jk}^2)\right] \tag{A.47}$$

with the individual helicity contributions

$$\bar{a}^{\mathrm{IF,gl}}_{g_+q_+\mapsto g_+g_+q_+} = \frac{1}{s_{AK}}\left[\frac{1}{y_{aj}y_{jk}} + \frac{1}{y_{aj}(1-y_{jk})} - \frac{\mu_k^2}{y_{jk}^2(1-y_{aj})}\right], \tag{A.48}$$

$$\bar{a}^{\mathrm{IF,gl}}_{g_+q_+\mapsto g_+g_-q_+} = \frac{1}{s_{AK}}\left[\frac{(1-y_{aj})^2 + [(1-y_{jk})^3 - 1](1-y_{aj})^2}{y_{aj}y_{jk}}\right. \tag{A.49}$$

$$\left. - \frac{\mu_k^2(1-y_{aj})(1-y_{jk})^3}{y_{jk}^2}\right], \tag{A.50}$$

$$\bar{a}^{\mathrm{IF,gl}}_{g_+q_+\mapsto g_-g_-q_+} = \frac{1}{s_{AK}}\frac{y_{jk}^3}{y_{aj}(1-y_{jk})}, \tag{A.51}$$

$$\bar{a}^{\mathrm{IF,gl}}_{g_+q_+\mapsto g_+g_+q_-} = \frac{1}{s_{AK}}\frac{\mu_k^2 y_{aj}^2}{y_{jk}^2(1-y_{aj})}, \tag{A.52}$$

JHEP07(2020)032

$$\bar{a}^{\mathrm{IF,gl}}_{g_+q_-\mapsto g_+g_+q_-} = \frac{1}{s_{AK}}\left[\frac{(1-y_{aj})^2}{y_{aj}y_{jk}} + \frac{1}{y_{aj}(1-y_{jk})} - \frac{\mu_k^2(1-y_{aj})}{y_{jk}^2}\right], \tag{A.50}$$

$$\bar{a}^{\mathrm{IF,gl}}_{g_+q_-\mapsto g_+g_-q_-} = \frac{1}{s_{AK}}\left[\frac{(1-y_{jk})^3}{y_{aj}y_{jk}} - \frac{\mu_k^2(1-y_{jk})^3}{y_{jk}^2(1-y_{aj})}\right]. \tag{A.51}$$

GGEmitIF. The helicity-averaged antenna function for the process $g_A g_K \mapsto g_a g_j g_k$ is

$$\bar{a}^{\mathrm{IF,gl}}_{g/gg}(y_{aj}, y_{jk}) = \frac{1}{s_{AK}}\left[\frac{(1-y_{aj})^3 + (1-y_{jk})^3}{y_{aj}y_{jk}} + \frac{1+y_{jk}^3}{y_{aj}(1-y_{jk})} + (1-\alpha)\frac{1-2y_{aj}}{y_{jk}} + 3 - 2y_{jk}\right] \tag{A.52}$$

with the helicity contributions

$$\bar{a}^{\mathrm{IF,gl}}_{g_+g_+\mapsto g_+g_+g_+} = \frac{1}{s_{AK}}\left[\frac{1}{y_{aj}y_{jk}} + (1-\alpha)\frac{1-2y_{aj}}{y_{jk}} + \frac{1}{y_{aj}(1-y_{jk})}\right], \tag{A.53}$$

$$\bar{a}^{\mathrm{IF,gl}}_{g_+g_+\mapsto g_+g_-g_+} = \frac{1}{s_{AK}}\left[\frac{(1-y_{aj})^3 + (1-y_{jk})^3 - 1}{y_{aj}y_{jk}} + 6 - 3y_{aj} - 3y_{jk} + y_{aj}y_{jk}\right], \tag{A.54}$$

$$\bar{a}^{\mathrm{IF,gl}}_{g_+g_-\mapsto g_+g_+g_-} = \frac{1}{s_{AK}}\left[\frac{(1-y_{aj})^3}{y_{aj}y_{jk}} + \frac{1}{y_{aj}(1-y_{jk})}\right], \tag{A.55}$$

$$\bar{a}^{\mathrm{IF,gl}}_{g_+g_-\mapsto g_+g_-g_-} = \frac{1}{s_{AK}}\left[\frac{(1-y_{jk})^3}{y_{aj}y_{jk}} + (1-\alpha)\frac{1-2y_{aj}}{y_{jk}} + 3y_{aj} - y_{jk} - y_{aj}y_{jk}\right]. \tag{A.56}$$

XGSplitIF. The helicity-averaged antenna function for the process $X_A g_K \mapsto X_a q_j \bar{q}_k$, where X is an arbitrary coloured parton in the initial state, is

$$\bar{a}^{\mathrm{IF,gl}}_{\bar{q}/Xg}(y_{aj}, y_{jk}) = \frac{1}{s_{AK}}\frac{y_{AK}}{2(y_{jk}+2\mu_j^2)}\left[y_{ak}^2 + y_{aj}^2 + \frac{2\mu_j^2}{y_{jk}+2\mu_j^2}\right]. \tag{A.57}$$

with the individual helicity contributions

$$\bar{a}^{\mathrm{IF,gl}}_{Xg_+\mapsto X\bar{q}_-q_+} = \frac{1}{s_{AK}}\frac{1}{2}\frac{y_{AK}}{y_{jk}+2\mu_j^2}\left[y_{ak}^2 - \frac{\mu_j^2 y_{ak}}{y_{jk}+2\mu_j^2(1-y_{ak})}\right], \tag{A.58}$$

$$\bar{a}^{\mathrm{IF,gl}}_{Xg_+\mapsto X\bar{q}_+q_-} = \frac{1}{s_{AK}}\frac{1}{2}\frac{y_{AK}}{y_{jk}+2\mu_j^2}\left[y_{aj}^2 - \frac{\mu_j^2 y_{aj}}{y_{jk}+2\mu_j^2(1-y_{aj})}\right], \tag{A.59}$$

$$\bar{a}^{\mathrm{IF,gl}}_{Xg_+\mapsto X\bar{q}_+q_+} = \frac{1}{s_{AK}}\frac{1}{2}\frac{\mu_j^2 y_{AK}}{(y_{jk}+2\mu_j^2)^2}\left[\frac{y_{aj}}{(1-y_{aj})} + \frac{y_{ak}}{(1-y_{ak})} + 2\right]. \tag{A.60}$$

QXSplitIF. The helicity-averaged antenna function for the process $q_A X_K \mapsto g_a \bar{q}_j X_k$, where X is an arbitrary coloured parton in the final state, is

$$\bar{a}^{\mathrm{IF,gl}}_{\bar{q}/qX}(y_{aj}, y_{jk}) = \frac{1}{s_{AK}}\left[\frac{y_{AK}^2 + (1-y_{AK})^2}{y_{aj}} + \frac{2\mu_j^2 y_{AK}}{y_{aj}^2}\right] \tag{A.61}$$

JHEP07(2020)032

with the individual helicity contributions

$$\bar{a}^{\mathrm{IF,gl}}_{q_+X\mapsto g_+\bar{q}_-X} = \frac{1}{s_{AK}}\left[\frac{y_{AK}^2}{y_{aj}} - \frac{\mu_j^2 y_{AK}^2}{y_{aj}^2(1-y_{AK})}\right], \tag{A.62}$$

$$\bar{a}^{\mathrm{IF,gl}}_{q_+X\mapsto g_-\bar{q}_-X} = \frac{1}{s_{AK}}\left[\frac{(1-y_{AK})^2}{y_{aj}} - \frac{\mu_j^2(1-y_{AK})}{y_{aj}^2}\right], \tag{A.63}$$

$$\bar{a}^{\mathrm{IF,gl}}_{q_+X\mapsto g_+\bar{q}_+X} = \frac{1}{s_{AK}}\frac{\mu_j^2}{y_{aj}^2(1-y_{AK})}. \tag{A.64}$$

GXConvIF. The helicity-averaged antenna function for the process $g_A X_K \mapsto q_a q_j X_k$, where X is an arbitrary coloured parton in the final state, is

$$\bar{a}^{\mathrm{IF,gl}}_{q/gX}(y_{aj}, y_{jk}) = \frac{1}{s_{AK}}\frac{1}{2}\left[\frac{1+(1-y_{AK})^2}{y_{AK}(y_{aj}-2\mu_j^2)} - \frac{2\mu_j^2 y_{AK}}{(y_{aj}-2\mu_j^2)^2}\right]. \tag{A.65}$$

with the individual helicity contributions

$$\bar{a}^{\mathrm{IF,gl}}_{g_+X\mapsto q_+q_+X} = \frac{1}{s_{AK}}\frac{1}{2}\left[\frac{1}{y_{AK}(y_{aj}-2\mu_j^2)} - \frac{\mu_j^2 y_{AK}}{(y_{aj}-2\mu_j^2)^2(1-y_{AK})}\right], \tag{A.66}$$

$$\bar{a}^{\mathrm{IF,gl}}_{g_+X\mapsto q_-q_-X} = \frac{1}{s_{AK}}\frac{1}{2}\left[\frac{(1-y_{AK})^2}{y_{AK}(y_{aj}-2\mu_j^2)} - \frac{\mu_j^2 y_{AK}(1-y_{AK})}{(y_{aj}-2\mu_j)^2}\right], \tag{A.67}$$

$$\bar{a}^{\mathrm{IF,gl}}_{g_+X\mapsto q_+q_-X} = \frac{1}{s_{AK}}\frac{1}{2}\frac{\mu_j^2}{(y_{aj}-2\mu_j^2)^2}\frac{y_{AK}^3}{1-y_{AK}}. \tag{A.68}$$

A.3 Initial-initial antennae

QQEmitII. The helicity-averaged antenna function for the process $\bar{q}_A q_B \mapsto q_a g_j q_b$ is

$$\bar{a}^{\mathrm{II,gl}}_{g/\bar{q}q}(y_{aj}, y_{jb}) = \frac{1}{s_{AB}}\left[\frac{(1-y_{aj})^2+(1-y_{jb})^2}{y_{aj}y_{jb}} + 1 - \frac{2\mu_a^2(1-y_{jb})}{y_{aj}^2} - \frac{2\mu_b^2(1-y_{aj})}{y_{jb}^2}\right] \tag{A.69}$$

with the individual helicity contributions

$$\bar{a}^{\mathrm{II,gl}}_{q_+q_+\mapsto q_+g_+q_+} = \frac{1}{s_{AB}}\left[\frac{1}{y_{aj}y_{jb}} - \frac{\mu_a^2}{y_{aj}^2} - \frac{\mu_b^2}{y_{jb}^2}\right], \tag{A.70}$$

$$\bar{a}^{\mathrm{II,gl}}_{q_+q_+\mapsto q_+g_-q_+} = \frac{1}{s_{AB}}\left[\frac{y_{AB}^2}{y_{aj}y_{jb}} - \frac{\mu_a^2(1-y_{jb})^2}{y_{aj}^2} - \frac{\mu_b^2(1-y_{aj})^2}{y_{jb}^2}\right], \tag{A.71}$$

$$\bar{a}^{\mathrm{II,gl}}_{q_+q_+\mapsto q_-g_-q_+} = \frac{1}{s_{AB}}\frac{\mu_a^2 y_{jb}^2}{y_{aj}^2}, \tag{A.72}$$

$$\bar{a}^{\mathrm{II,gl}}_{q_+q_+\mapsto q_+g_-q_-} = \frac{1}{s_{AB}}\frac{\mu_b^2 y_{aj}^2}{y_{jb}^2}, \tag{A.73}$$

JHEP07(2020)032

$$\bar{a}^{\mathrm{II,gl}}_{q_+q_-\mapsto q_+g_+q_-} = \frac{1}{s_{AB}}\left[\frac{(1-y_{aj})^2}{y_{aj}y_{jb}} - \frac{\mu_a^2}{y_{aj}^2} - \frac{\mu_b^2(1-y_{aj})^2}{y_{jb}^2}\right], \tag{A.74}$$

$$\bar{a}^{\mathrm{II,gl}}_{q_+q_-\mapsto q_+g_-q_-} = \frac{1}{s_{AB}}\left[\frac{(1-y_{jb})^2}{y_{aj}y_{jb}} - \frac{\mu_a^2(1-y_{jb})^2}{y_{aj}^2} - \frac{\mu_b^2}{y_{jb}^2}\right], \tag{A.75}$$

$$\bar{a}^{\mathrm{II,gl}}_{q_+q_-\mapsto q_-g_-q_-} = \frac{1}{s_{AB}}\frac{\mu_a^2 y_{jb}^2}{y_{aj}^2}, \tag{A.76}$$

$$\bar{a}^{\mathrm{II,gl}}_{q_+q_-\mapsto q_-g_+q_+} = \frac{1}{s_{AB}}\frac{\mu_b^2 y_{aj}^2}{y_{jb}^2}. \tag{A.77}$$

QGEmitII. The helicity-averaged antenna function for the process $\bar{q}_A g_B \mapsto q_a g_j g_b$ is

$$\bar{a}^{\mathrm{II,gl}}_{g/qg}(y_{aj}, y_{jb}) = \frac{1}{s_{AB}}\left[\frac{(1-y_{aj})^3 + (1-y_{jb})^2}{y_{aj}y_{jb}} + \frac{1+y_{aj}^3}{y_{jb}(1-y_{aj})} - \frac{2\mu_a^2(1-y_{jb})}{y_{aj}^2} + 2 - y_{aj} - \frac{y_{jb}}{2}\right] \tag{A.78}$$

with the individual helicity contributions

$$\bar{a}^{\mathrm{II,gl}}_{q_+g_+\mapsto q_+g_+g_+} = \frac{1}{s_{AB}}\left[\frac{1}{y_{aj}y_{jb}} + \frac{1}{y_{jb}(1-y_{aj})} - \frac{\mu_a^2}{y_{aj}^2}\right]. \tag{A.79}$$

$$\bar{a}^{\mathrm{II,gl}}_{q_+g_+\mapsto q_+g_-g_+} = \frac{1}{s_{AB}}\left[\frac{(1-y_{aj})y_{AB}^2}{y_{aj}y_{jb}} - \frac{\mu_a^2(1-y_{jb})^2}{y_{aj}^2}\right], \tag{A.80}$$

$$\bar{a}^{\mathrm{II,gl}}_{q_+g_+\mapsto q_-g_-g_+} = \frac{1}{s_{AB}}\left[\frac{\mu_a^2 y_{jb}^2}{y_{aj}^2}\right], \tag{A.81}$$

$$\bar{a}^{\mathrm{II,gl}}_{q_+g_+\mapsto q_+g_-g_-} = \frac{1}{s_{AB}}\frac{y_{aj}^3}{y_{jb}(1-y_{aj})}, \tag{A.82}$$

$$\bar{a}^{\mathrm{II,gl}}_{q_+g_-\mapsto q_+g_+g_-} = \frac{1}{s_{AB}}\left[\frac{(1-y_{aj})^3}{y_{aj}y_{jb}} - \frac{\mu_a^2}{y_{aj}^2}\right], \tag{A.83}$$

$$\bar{a}^{\mathrm{II,gl}}_{q_+g_-\mapsto q_+g_-g_-} = \frac{1}{s_{AB}}\left[\frac{(1-y_{jb})^2}{y_{aj}y_{jb}} + \frac{1}{y_{jb}(1-y_{aj})} - \frac{\mu_a^2(1-y_{jb})^2}{y_{aj}^2}\right]. \tag{A.84}$$

GGEmitII. The helicity-averaged antenna function for the process $g_A g_B \mapsto g_a g_j g_b$ is

$$\bar{a}^{\mathrm{II,gl}}_{g/gg}(y_{aj}, y_{jb}) = \frac{1}{s_{AB}}\left[\frac{(1-y_{aj})^3 + (1-y_{jb})^3}{y_{aj}y_{jb}} + \frac{1+y_{aj}^3}{y_{jb}(1-y_{aj})} + \frac{1+y_{jb}^3}{y_{aj}(1-y_{jb})} + 3 - \frac{3y_{aj}}{2} - \frac{3y_{jb}}{2}\right] \tag{A.85}$$

with the individual helicity contributions

$$\bar{a}^{\mathrm{II,gl}}_{g_+g_+\mapsto g_+g_+g_+} = \frac{1}{s_{AB}}\left[\frac{1}{y_{aj}y_{jb}} + \frac{1}{y_{jb}(1-y_{aj})} + \frac{1}{y_{aj}(1-y_{jb})}\right], \tag{A.86}$$

$$\bar{a}^{\mathrm{II,gl}}_{g_+g_+\mapsto g_+g_-g_+} = \frac{1}{s_{AB}}\frac{y_{AB}^3}{y_{aj}y_{jb}}, \tag{A.87}$$

JHEP07(2020)032

$$\bar{a}^{\mathrm{II,gl}}_{g_+g_-\mapsto g_+g_+g_-} = \frac{1}{s_{AB}}\left[\frac{(1-y_{aj})^3}{y_{aj}y_{jb}} + \frac{1}{y_{aj}(1-y_{jb})}\right], \tag{A.88}$$

$$\bar{a}^{\mathrm{II,gl}}_{g_+g_-\mapsto g_+g_-g_-} = \frac{1}{s_{AB}}\left[\frac{(1-y_{jb})^3}{y_{aj}y_{jb}} + \frac{1}{y_{jb}(1-y_{aj})}\right], \tag{A.89}$$

$$\bar{a}^{\mathrm{II,gl}}_{g_+g_+\mapsto g_+g_-g_-} = \frac{1}{s_{AB}}\frac{y_{aj}^3}{y_{jb}(1-y_{aj})}, \tag{A.90}$$

$$\bar{a}^{\mathrm{II,gl}}_{g_+g_+\mapsto g_-g_-g_+} = \frac{1}{s_{AB}}\frac{y_{jb}^3}{y_{aj}(1-y_{jb})}. \tag{A.91}$$

QXSplitII. The helicity-averaged antenna function for the process $q_A X_B \mapsto g_a \bar{q}_j X_b$, where X is an arbitrary coloured parton, is

$$\bar{a}^{\mathrm{II,gl}}_{\bar{q}/qX}(y_{aj}, y_{jb}) = \frac{1}{s_{AB}}\left[\frac{y_{AB}^2 + (1-y_{AB})^2}{y_{aj}} + \frac{2\mu_j^2 y_{AB}}{y_{aj}^2}\right] \tag{A.92}$$

with the individual helicity contributions

$$\bar{a}^{\mathrm{II,gl}}_{q_+X\mapsto g_+\bar{q}_-X} = \frac{1}{s_{AB}}\left[\frac{y_{AB}^2}{y_{aj}} - \frac{\mu_j^2 y_{AB}^2}{y_{aj}^2(1-y_{AB})}\right], \tag{A.93}$$

$$\bar{a}^{\mathrm{II,gl}}_{q_+X\mapsto g_-\bar{q}_-X} = \frac{1}{s_{AB}}\left[\frac{(1-y_{AB})^2}{y_{aj}} - \frac{\mu_j^2(1-y_{AB})}{y_{aj}^2}\right], \tag{A.94}$$

$$\bar{a}^{\mathrm{II,gl}}_{q_+X\mapsto g_+\bar{q}_+X} = \frac{1}{s_{AB}}\frac{\mu_j^2}{y_{aj}^2(1-y_{AB})}. \tag{A.95}$$

GXConvII. The helicity-averaged antenna function for the process $g_A X_B \mapsto q_a q_j X_b$, where X is an arbitrary coloured parton, is

$$\bar{a}^{\mathrm{II,gl}}_{q/gX}(s_{AB}, y_{aj}, y_{jb}) = \frac{1}{s_{AB}}\left[\frac{1+(1-y_{AB})^2}{(y_{aj}-2\mu_j^2)y_{AB}} - \frac{2\mu_j^2 y_{AB}}{(y_{aj}-2\mu_j^2)^2}\right]. \tag{A.96}$$

with the individual helicity contributions

$$\bar{a}^{\mathrm{II,gl}}_{g_+X\mapsto q_+q_+X} = \frac{1}{s_{AB}}\frac{1}{2}\left[\frac{1}{(y_{aj}-2\mu_j^2)y_{AB}} - \frac{\mu_j^2}{(y_{aj}-2\mu_j^2)^2}\frac{y_{AB}}{1-y_{AB}}\right], \tag{A.97}$$

$$\bar{a}^{\mathrm{II,gl}}_{g_+X\mapsto q_-q_-X} = \frac{1}{s_{AB}}\left[\frac{(1-y_{AB})^2}{(y_{aj}-2\mu_j^2)y_{AB}} - \frac{\mu_j^2 y_{AB}(1-y_{AB})}{(y_{aj}-2\mu_j^2)^2}\right], \tag{A.98}$$

$$\bar{a}^{\mathrm{II,gl}}_{g_+X\mapsto q_+q_-X} = \frac{1}{s_{AB}}\frac{\mu_j^2}{(y_{aj}-2\mu_j^2)^2}\frac{y_{AB}^3}{1-y_{AB}}. \tag{A.99}$$

B Collinear limits of sector antennae

To ensure that the sector (as well as the global) antenna functions reproduce the correct collinear limits, they are numerically tested against DGLAP splitting kernels. For two final-state partons (i.e. in final-final or initial-final antennae), we compare antenna

JHEP07(2020)032

functions against massive DGLAP kernels in the quasi-collinear limit, while for initial-final and initial-initial configurations, we compare them to massless DGLAP kernels in the collinear limit.

Below, we collect all helicity- and mass-dependent DGLAP kernels as implemented in VINCIA [84], denoting them by $P_{I_\pm \mapsto i_\pm j_\pm}(z)$, where $z \equiv z_i$ is the energy fraction taken by parton i.

The helicity-averaged splitting kernel for gluon emissions $g \mapsto gg$ is given by

$$P_{g\mapsto gg}(z) = 2\frac{(1-z(1-z))^2}{z(1-z)} \tag{B.1}$$

and the individual helicity contributions are

$$P_{g_+\mapsto g_+g_+}(z) = \frac{1}{z(1-z)}, \tag{B.2}$$

$$P_{g_+\mapsto g_-g_+}(z) = \frac{(1-z)^3}{z}, \tag{B.3}$$

$$P_{g_+\mapsto g_+g_-}(z) = \frac{z^3}{1-z}. \tag{B.4}$$

For gluon emissions from a quark, $q \mapsto qg$, the helicity-averaged splitting kernel is

$$P_{q\mapsto qg}(z) = \frac{1+z^2}{1-z} - 2\mu_q^2 \tag{B.5}$$

and the individual helicity contributions are

$$P_{q_+\mapsto q_+g_+}(z) = \frac{1}{1-z} - \mu_q^2\frac{1}{z}, \tag{B.6}$$

$$P_{q_+\mapsto q_-g_+}(z) = \mu_q^2\frac{(1-z)^2}{z}, \tag{B.7}$$

$$P_{q_+\mapsto q_+g_-}(z) = \frac{z^2}{1-z} - \mu_q^2 z. \tag{B.8}$$

Lastly, for gluon splittings $g \mapsto q\bar{q}$, the helicity-averaged splitting kernel is

$$P_{g\mapsto q\bar{q}}(z) = z^2 + (1-z)^2 + 2\mu_q^2 \tag{B.9}$$

and the individual helicity contributions are given by

$$P_{g_+\mapsto q_+\bar{q}_-}(z) = z^2 - \mu_q^2\frac{z}{1-z}, \tag{B.10}$$

$$P_{g_+\mapsto q_-\bar{q}_+}(z) = (1-z)^2 - \mu_q^2\frac{1-z}{z}, \tag{B.11}$$

$$P_{g_+\mapsto q_+\bar{q}_+}(z) = \mu_q^2\left(\frac{z}{1-z} + \frac{1-z}{z} + 2\right). \tag{B.12}$$

C Explicit expressions for trial generators

Here, we collect all additional trial integrals, eq. (3.10), needed for the sector shower in terms of the phase space variables $x_\perp$ and ζ. We denote trial functions used only in the sector shower by a superscript "sct" and express trial integrals in terms of eq. (3.20).

JHEP07(2020)032

C.1 Trial integrals

FF branchings. For final-final configurations, the trial integral we are inverting analytically is given by

$$\mathcal{A}(x_{\perp 1}^{\mathrm{FF}}, x_{\perp 2}^{\mathrm{FF}}) = f_{\mathrm{Källén}}^{\mathrm{FF}} \frac{\mathcal{C}}{4\pi} \int_{x_{\perp 2}^{\mathrm{FF}}}^{x_{\perp 1}^{\mathrm{FF}}} \alpha_{\mathrm{s}}(x_{\perp}^{\mathrm{FF}}) \bar{a}_{\mathrm{trial}}(x_{\perp}^{\mathrm{FF}}, \zeta^{\mathrm{FF}}) s_{IK} |J(x_{\perp}^{\mathrm{FF}}, \zeta^{\mathrm{FF}})| \mathrm{d}x_{\perp}^{\mathrm{FF}} \mathrm{d}\zeta^{\mathrm{FF}}. \quad \text{(C.1)}$$

To ensure that the trial function overestimates the sector antenna function, we multiply the gluon splitting trial function,

$$\bar{a}_{\text{trial-split}}^{\mathrm{FF}} = \frac{1}{s_{IK}} \frac{1}{2(y_{ij} + 2\mu_j^2)}, \quad \text{(C.2)}$$

by two to take into account that the splitting gluon is part of two antennae and overestimate the additional collinear parts of gluon-emission sector antenna functions in eq. (3.3) by

$$\bar{a}_{\text{trial-coll-}I}^{\mathrm{FF,sct}} = \frac{1}{s_{IK}} \frac{2}{y_{ij}(1 - y_{jk})}, \quad \text{(C.3)}$$

$$\bar{a}_{\text{trial-coll-}K}^{\mathrm{FF,sct}} = \frac{1}{s_{IK}} \frac{2}{y_{jk}(1 - y_{ij})}. \quad \text{(C.4)}$$

For these additional collinear trial functions, both in gluon emissions and gluon splittings, we choose the corresponding non-singular invariant,

$$\zeta^{\mathrm{FF}} = \begin{cases} \zeta_-^{\mathrm{FF}} = y_{jk} + \mu_j^2 & I\text{-collinear} \\ \zeta_+^{\mathrm{FF}} = y_{ij} + \mu_j^2 & K\text{-collinear} \end{cases}, \quad \text{(C.5)}$$

where $\mu_j^2 = 0$ for gluon emissions. The Jacobian to transform from (y_{ij}, y_{jk}) to $(x_{\perp}^{\mathrm{FF}}, \zeta^{\mathrm{FF}})$ is given by

$$|J| = |J_{\pm}| = \frac{1}{\zeta_{\pm}^{\mathrm{FF}}}. \quad \text{(C.6)}$$

In terms of eq. (3.20), the additional trial integrals are given by

$$\mathcal{A}_{\mathrm{coll}}\left(x_{\perp 1}^{\mathrm{FF}}, x_{\perp 2}^{\mathrm{FF}}\right) = f_{\mathrm{Källén}}^{\mathrm{FF}} \frac{\mathcal{C}}{2\pi} \bar{\mathcal{A}}\left(x_{\perp 1}^{\mathrm{FF}}, x_{\perp 2}^{\mathrm{FF}}, \frac{1}{1 - \zeta_{\pm}^{\mathrm{FF}}}\right), \quad \text{(C.7)}$$

$$\mathcal{A}_{\mathrm{split}}\left(x_{\perp 1}^{\mathrm{FF}}, x_{\perp 2}^{\mathrm{FF}}\right) = f_{\mathrm{Källén}}^{\mathrm{FF}} \frac{\mathcal{C}}{2\pi} \bar{\mathcal{A}}\left(x_{\perp 1}^{\mathrm{FF}}, x_{\perp 2}^{\mathrm{FF}}, \frac{1}{2}\right), \quad \text{(C.8)}$$

which are the same for the I-collinear and K-collinear regions, and the choice of $\zeta_{\pm}^{\mathrm{FF}}$ as the corresponding non-singular invariant is implicit.

RF branchings. For resonance-final configurations, the trial integral is given by:

$$\mathcal{A}(x_{\perp 1}^{\mathrm{RF}}, x_{\perp 2}^{\mathrm{RF}}) = f_{\mathrm{Källén}}^{\mathrm{RF}} \frac{\mathcal{C}}{4\pi} \int_{x_{\perp 2}^{\mathrm{RF}}}^{x_{\perp 1}^{\mathrm{RF}}} \alpha_{\mathrm{s}}(x_{\perp}^{\mathrm{RF}}) \bar{a}_{\mathrm{trial}}(x_{\perp}^{\mathrm{RF}}, \zeta^{\mathrm{RF}}) \frac{(s_{AK} + m_j^2 + m_k^2 - m_K^2)^2}{(1 - y_{jk})^3} \times |J(x_{\perp}^{\mathrm{RF}}, \zeta^{\mathrm{RF}})| \mathrm{d}x_{\perp}^{\mathrm{RF}} \mathrm{d}\zeta^{\mathrm{RF}}. \quad \text{(C.9)}$$

JHEP07(2020)032

The trial antenna for (K-) gluon splittings is given by

$$\bar{a}^{\text{RF,sct}}_{\text{trial-split-}K} = \left[\frac{1-y_{jk}}{s_{AK}+2m_j^2}\right]\frac{2\,(1-y_{jk})^2}{(y_{jk}+2\mu_j^2)}\,. \tag{C.10}$$

The first factor in square brackets simply results from the dimensionful normalisation of the trial antenna.

The additional trial antenna for gluon emission collinear with parton K is given by:

$$\bar{a}^{\text{RF,sct}}_{\text{trial-coll-}K} = \left[\frac{1}{s_{AK}}\right]\frac{2}{y_{jk}(1-y_{aj})}\left[2(1-y_{jk})\right]^3\,. \tag{C.11}$$

The final term inside the brackets is simply $y_{AK}/y_{AK\,\text{min}} \geq 1$, and is included to cancel the factor appearing in eq. (2.30). We also note that since A is a resonance, and therefore cannot be a gluon, there is no additional A-collinear piece present in the antenna function and hence there is no need for the corresponding overestimate.

We make the following choice for the complementary phase space variable[13]

$$\zeta^{\text{RF}} = \zeta_+^{\text{RF}} = y_{aj}\,, \tag{C.12}$$

leading to the Jacobian

$$|J| = |J_+| = \frac{1}{y_{aj}-\mu_j^2}\,, \tag{C.13}$$

where $\mu_j^2 = 0$ for gluon emissions.

In terms of eq. (3.20), the sector trial integrals are given by

$$\mathcal{A}_{\text{coll-}K}\left(x_{\perp 1}^{\text{RF}}, x_{\perp 2}^{\text{RF}}\right) = f^{\text{RF}}_{\text{Källén}}\frac{\mathcal{C}}{2\pi}\bar{\mathcal{A}}\left(x_{\perp 1}^{\text{FF}}, x_{\perp 2}^{\text{FF}}, \frac{8}{1-\zeta_+^{\text{RF}}}\right) \tag{C.14}$$

$$\mathcal{A}_{\text{split}}\left(x_{\perp 1}^{\text{RF}}, x_{\perp 2}^{\text{RF}}\right) = f^{\text{RF}}_{\text{Källén}}\frac{\mathcal{C}}{2\pi}\bar{\mathcal{A}}\left(x_{\perp 1}^{\text{FF}}, x_{\perp 2}^{\text{FF}}, 1\right). \tag{C.15}$$

IF branchings. For initial-final configurations, the trial integral is

$$\mathcal{A}(x_{\perp 1}, x_{\perp 2}) = \frac{\mathcal{C}}{4\pi}\tilde{R}_f\int_{x_{\perp 2}}^{x_{\perp 1}}\alpha_{\text{s}}(x_\perp^{\text{IF}})\bar{a}_{\text{trial}}(x_\perp^{\text{IF}}, \zeta^{\text{IF}})y_{AK}^{\gamma}\frac{s_{AK}}{1-y_{jk}}|J(x_\perp^{\text{IF}}, \zeta^{\text{IF}})|\mathrm{d}x_\perp^{\text{IF}}\mathrm{d}\zeta^{\text{IF}} \tag{C.16}$$

as $x_A/x_a = y_{AK}$, cf. section 2.3.

Similarly to the case of final-final splittings, we overestimate gluon splittings in the final state by

$$\bar{a}^{\text{IF,sct}}_{\text{trial-split-}K} = 2\cdot\frac{1}{s_{AK}}\frac{1}{2(y_{jk}+2\mu_j^2)}\left[\frac{s_{AK}+m_j^2}{s_{AK}}\right], \tag{C.17}$$

where the factor of 2 again takes into account that the gluon is part of two antennae and the additional factor in the brackets, which is ≥ 1, cancels unwanted factors in the trial integral. Gluon splittings and gluon conversions in the initial state are already sectorised

[13] We note that this choice is not completely identical to that made in [50] for gluon splittings; however, this was a more natural choice for the normalisation chosen in eq. (2.8) and the only impact of this choice is in the efficiency of the exploration of phase space, expected to be minimal.

JHEP07(2020)032

and do not need additional sector overestimates. In the sector shower, we complement this by another overestimate of the K-collinear part which is added for sector antennae. Hence, the only additional gluon-emission trial functions is

$$\bar{a}^{\text{IF,sct}}_{\text{trial-coll-}K} = \frac{1}{s_{AK}} \frac{2}{y_{jk}(1-y_{aj})} \,. \tag{C.18}$$

We emphasise that although the additional singularity $y_{aj} \to 1$ is kinematically allowed, it is regularised by the parton shower cutoff Q^2_{cut}, as it lies in the a-k (soft-)collinear sector. Hence, the sector-shower phase space limit

$$y^{\max}_{aj} = \frac{s_{aj}}{s_{aj} + s^{\min}_{ak}} = \frac{1}{1 + x_A \frac{Q^2_{\text{cut}}}{s_{AK}}} \tag{C.19}$$

is imposed in the shower evolution.

We use the K-collinear phase space variable

$$\zeta^{\text{IF}} = \zeta^{\text{IF}}_{+} = y_{aj} - \mu^2_j \,, \tag{C.20}$$

with the Jacobian for the transformation $(y_{aj}, y_{jk}) \mapsto (x^{\text{IF}}_{\perp}, \zeta^{\text{IF}})$

$$|J| = |J_+| = \frac{1}{\zeta^{\text{IF}}_{+}} \,. \tag{C.21}$$

For final-state gluon splittings, we therefore find the sector trial integral

$$\mathcal{A}_{\text{split-}K}(x^{\text{IF}}_{\perp 1}, x^{\text{IF}}_{\perp 2}) = \frac{\mathcal{C}}{2\pi} \tilde{R}_f \bar{\mathcal{A}} \left(x^{\text{IF}}_{\perp 1}, x^{\text{IF}}_{\perp 2}, \frac{1}{2} \right) \,, \tag{C.22}$$

while for gluon emissions, we have the sector trial integral

$$\mathcal{A}_{\text{coll-}K}(x^{\text{IF}}_{\perp 1}, x^{\text{IF}}_{\perp 2}) = \frac{\mathcal{C}}{2\pi} \tilde{R}_f \bar{\mathcal{A}} \left(x^{\text{IF}}_{\perp 1}, x^{\text{IF}}_{\perp 2}, \frac{1}{1-\zeta^{\text{IF}}_{+}} \right) \,, \tag{C.23}$$

again expressed in terms of eq. (3.20).

It should be pointed out that the additional sector integrals eqs. (C.22) and (C.23) are given only for $\gamma = 1$. In practice, γ is chosen to be either 0 or 1 for the others.

II branchings. No additional trial generators are needed for the initial-initial case, because global antenna functions are sectorised in the first place, as there is no emission into the initial state. The same trial integrals are used for the global and the sector shower.

There is, however, the possibility to express these trial integrals in terms of the dimensionless phase space variables $x^{\text{II}}_{\perp}$ and ζ^{II}. The trial integral is then given by

$$\mathcal{A}(x^{\text{II}}_{\perp 1}, x^{\text{II}}_{\perp 2}) = \frac{\mathcal{C}}{4\pi} \tilde{R}_f \int_{x^{\text{II}}_{\perp 2}}^{x^{\text{II}}_{\perp 1}} \alpha_{\text{s}}(x^{\text{II}}_{\perp}) \bar{a}_{\text{trial}}(x^{\text{II}}_{\perp}, \zeta^{\text{II}}) y^{\gamma}_{AB} \frac{s_{AB}}{1 - y_{aj} - y_{jb}} |J(x^{\text{II}}_{\perp}, \zeta^{\text{II}})| \mathrm{d}x^{\text{II}}_{\perp} \mathrm{d}\zeta^{\text{II}} \tag{C.24}$$

as $x_A x_B / x_a x_b = y_{AB}$, cf. section 2.3.

JHEP07(2020)032

C.2 Zeta integrals

Apart from the trivial integral

$$I_{\text{lin}}(\zeta) = \int 1 \mathrm{d}\zeta = \zeta \tag{C.25}$$

we make use of the integral

$$I_{\log}(\zeta) = \int \frac{1}{1-\zeta} \mathrm{d}\zeta = -\log(1-\zeta)\,. \tag{C.26}$$

Trial values for ζ are then found by inverting eq. (3.18) with the solution

$$\zeta = I_\zeta^{-1}\left[R_\zeta\left(I(\zeta_{\max}) - I(\zeta_{\min})\right) + I(\zeta_{\min})\right]\,, \tag{C.27}$$

where I_ζ^{-1} denotes the inverse of I_ζ, given by

$$I_{\text{lin}}^{-1}(r) = r\,, \tag{C.28}$$

$$I_{\log}^{-1}(r) = 1 - \mathrm{e}^{-r}\,. \tag{C.29}$$

JHEP07(2020)032

D Comparisons to high-multiplicity matrix elements

Here, we present validations of the sector shower against high-multiplicity tree-level matrix elements. For the sake of clarity, we do not include separate sector histograms, as too many histories contribute to each final state, although for a given kinematic configuration only a single one contributes at a time.

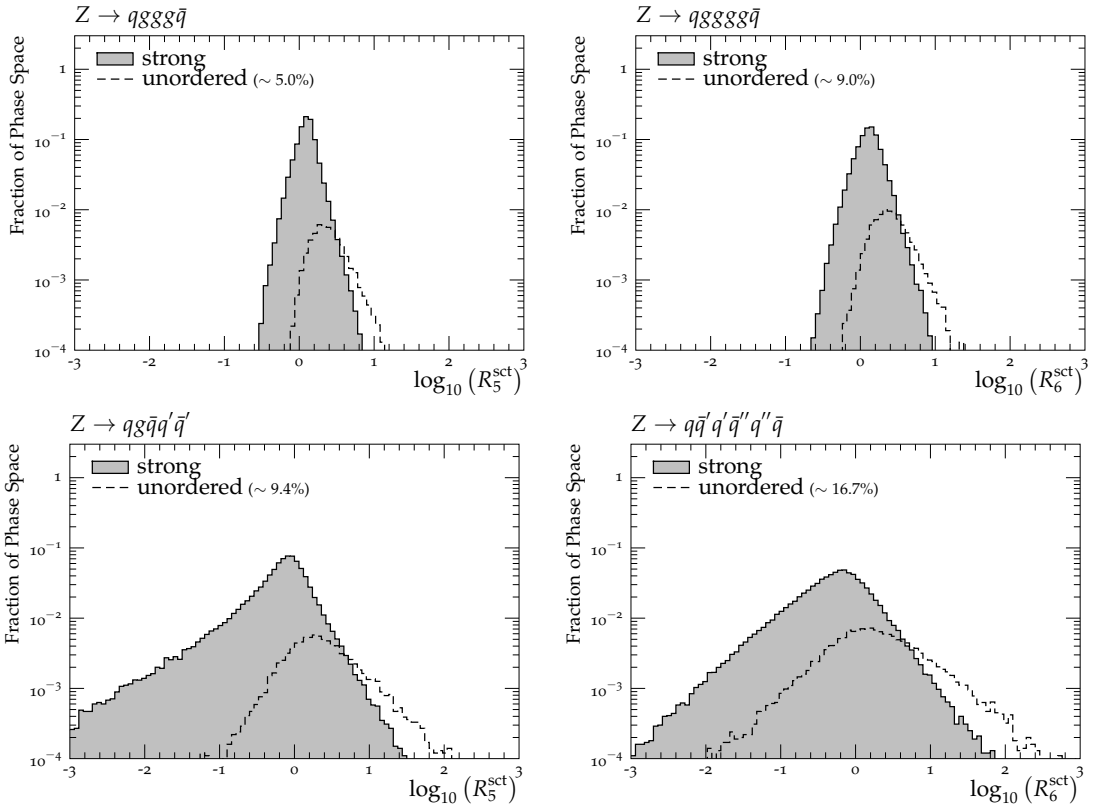


Figure 15. Ratios R_5^{sct} of the sector shower approximation to LO matrix elements for Z decays to 5- and 6- parton final states in a flat phase space scan.

JHEP07(2020)032

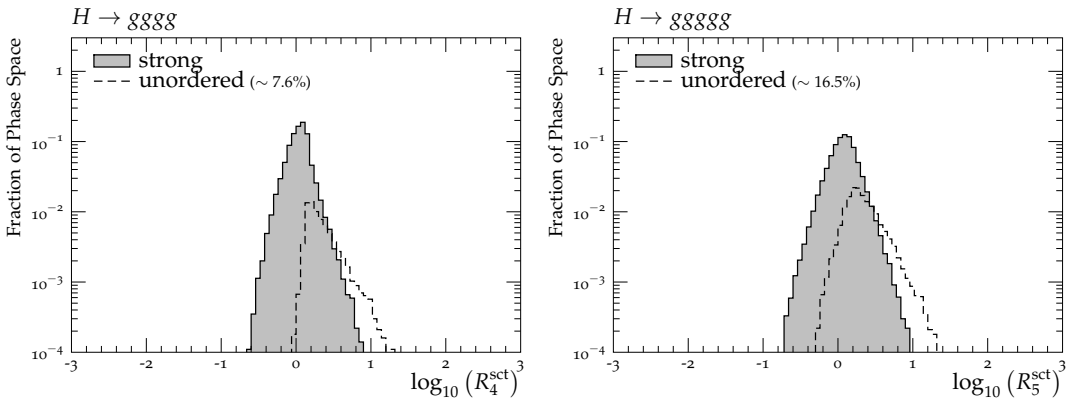


Figure 16. Ratios R_n^{sct} of the sector-shower approximation to LO matrix elements for $H \to ng$ with $n = 4, 5$ in a flat phase space scan.

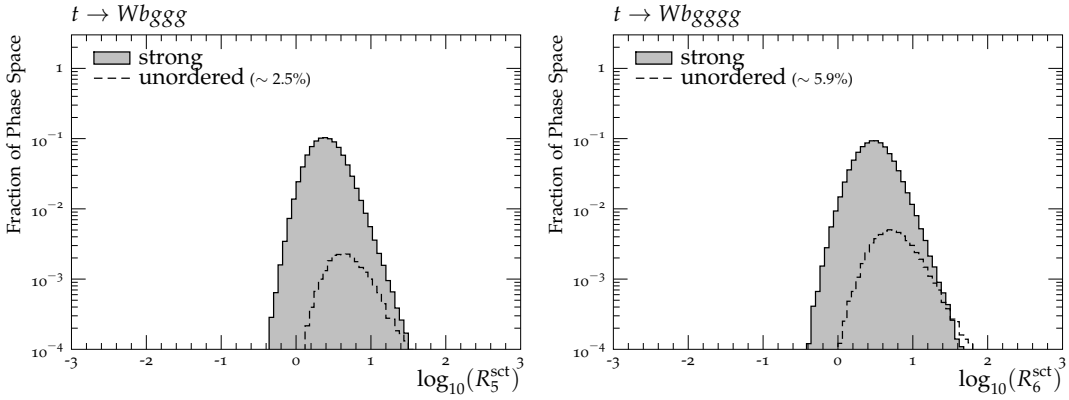


Figure 17. Ratios R_n^{sct} of the sector-shower approximation to LO matrix elements for top-quark decay processes $t \to W^+b$ with additional gluons in a flat phase space scan.

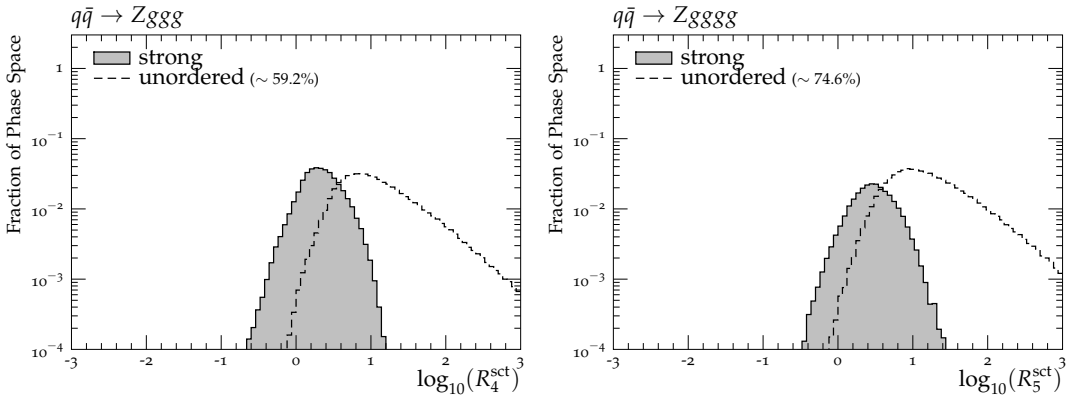


Figure 18. Ratios R_n^{sct} of the sector-shower approximation to LO matrix elements for Drell-Yan processes $q\bar{q} \to Z$ with additional gluons in a flat phase space scan.

JHEP07(2020)032

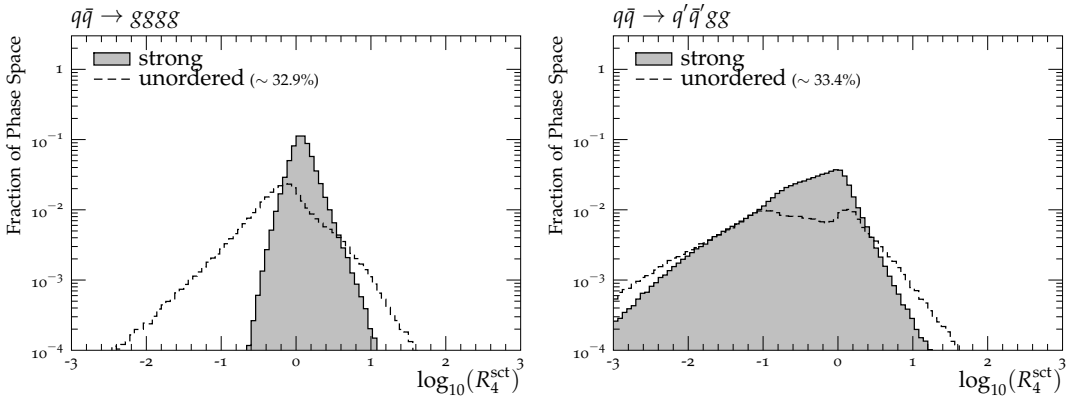


Figure 19. Ratios R_n^{sct} of the sector-shower approximation to LO matrix elements for $q\bar{q}$-annihilation processes to 4 partons in a flat phase space scan.

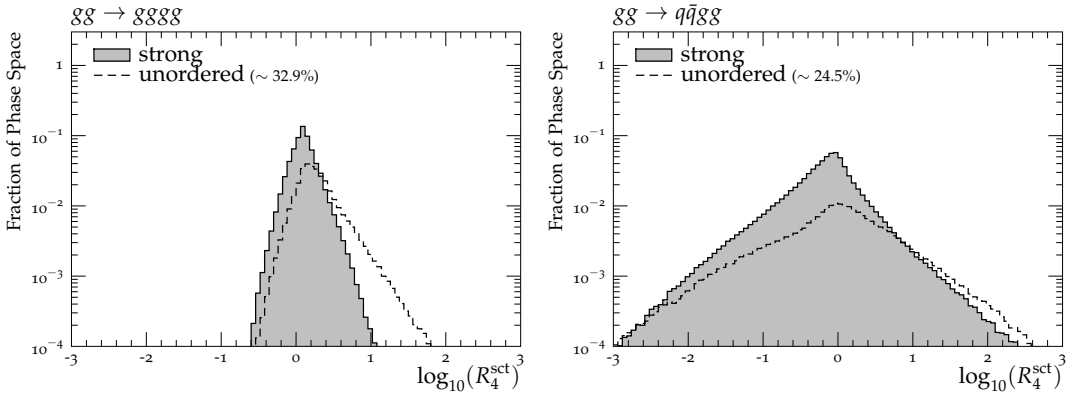


Figure 20. Ratios R_n^{sct} of the sector-shower approximation to LO matrix elements for gg-fusion processes to 4 partons in a flat phase space scan.

JHEP07(2020)032

E Tune parameters

The full set of parameters changed relative to the PYTHIA 8.3 default tune are shown in the following table.

Name	Now	Default
BeamRemnants:primordialKThard	0.4	1.8
BeamRemnants:primordialKTsoft	0.25	0.90
ColourReconnection:range	1.75	1.80
Diffraction:mMinPert	1000000.0	10.0
MultipartonInteractions:alphaSorder	2	1
MultipartonInteractions:alphaSvalue	0.119	0.130
MultipartonInteractions:ecmPow	0.210	0.215
MultipartonInteractions:expPow	1.75	1.85
MultipartonInteractions:pT0Ref	2.24	2.28
SigmaProcess:alphaSorder	2	1
SigmaProcess:alphaSvalue	0.119	0.130
StringFlav:etaPrimeSup	0.10	0.12
StringFlav:etaSup	0.5	0.6
StringFlav:mesonCvector	1.30	0.88
StringFlav:mesonSvector	0.53	0.55
StringFlav:mesonUDvector	0.42	0.50
StringFlav:popcornSmeson	0.75	0.50
StringFlav:popcornSpair	0.75	0.90
StringFlav:probQQ1toQQ0	0.025	0.0275
StringFlav:probQQtoQ	0.077	0.081
StringFlav:probSQtoQQ	1.000	0.915
StringFlav:probStoUD	0.205	0.217
StringPT:sigma	0.305	0.335
StringZ:aExtraDiquark	0.90	0.97
StringZ:aLund	0.55	0.68
StringZ:bLund	0.78	0.98
StringZ:rFactB	0.850	0.855
StringZ:rFactC	1.15	1.32

(E.1)

References

[1] J. Bellm et al., *HERWIG 7.0/HERWIG++ 3.0 release note*, *Eur. Phys. J. C* **76** (2016) 196 [arXiv:1512.01178] [INSPIRE].

[2] T. Gleisberg et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007 [arXiv:0811.4622] [INSPIRE].

JHEP07(2020)032

[3] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [arXiv:1410.3012] [INSPIRE].

[4] V.N. Gribov and L.N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438 [INSPIRE].

[5] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, *Nucl. Phys. B* **126** (1977) 298 [INSPIRE].

[6] Y.L. Dokshitzer, *Calculation of the structure functions for deep inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics*, *Sov. Phys. JETP* **46** (1977) 641 [INSPIRE].

[7] G. Marchesini and B.R. Webber, *Simulation of QCD jets including soft gluon interference*, *Nucl. Phys. B* **238** (1984) 1 [INSPIRE].

[8] M. Bengtsson and T. Sjöstrand, *A comparative study of coherent and noncoherent parton shower evolution*, *Nucl. Phys. B* **289** (1987) 810 [INSPIRE].

[9] G. Gustafson and U. Pettersson, *Dipole formulation of QCD cascades*, *Nucl. Phys. B* **306** (1988) 746 [INSPIRE].

[10] L. Lönnblad, *ARIADNE version 4: a program for simulation of QCD cascades implementing the color dipole model*, *Comput. Phys. Commun.* **71** (1992) 15 [INSPIRE].

[11] W.T. Giele, D.A. Kosower and P.Z. Skands, *A simple shower and matching algorithm*, *Phys. Rev. D* **78** (2008) 014026 [arXiv:0707.3652] [INSPIRE].

[12] W.T. Giele, D.A. Kosower and P.Z. Skands, *Higher-order corrections to timelike jets*, *Phys. Rev. D* **84** (2011) 054003 [arXiv:1102.2126] [INSPIRE].

[13] S. Catani and M.H. Seymour, *The dipole formalism for the calculation of QCD jet cross-sections at next-to-leading order*, *Phys. Lett. B* **378** (1996) 287 [hep-ph/9602277] [INSPIRE].

[14] S. Catani and M.H. Seymour, *A general algorithm for calculating jet cross-sections in NLO QCD*, *Nucl. Phys. B* **485** (1997) 291 [*Erratum ibid.* **510** (1998) 503] [hep-ph/9605323] [INSPIRE].

[15] Z. Nagy and D.E. Soper, *Matching parton showers to NLO computations*, *JHEP* **10** (2005) 024 [hep-ph/0503053] [INSPIRE].

[16] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani-Seymour dipole factorisation*, *JHEP* **03** (2008) 038 [arXiv:0709.1027] [INSPIRE].

[17] J.-C. Winter and F. Krauss, *Initial-state showering based on colour dipoles connected to incoming parton lines*, *JHEP* **07** (2008) 040 [arXiv:0712.3913] [INSPIRE].

[18] S. Platzer and S. Gieseke, *Coherent parton showers with local recoils*, *JHEP* **01** (2011) 024 [arXiv:0909.5593] [INSPIRE].

[19] S. Höche and S. Prestel, *The midpoint between dipole and parton showers*, *Eur. Phys. J. C* **75** (2015) 461 [arXiv:1506.05057] [INSPIRE].

[20] A. Gehrmann-De Ridder, M. Ritzmann and P.Z. Skands, *Timelike dipole-antenna showers with massive fermions*, *Phys. Rev. D* **85** (2012) 014013 [arXiv:1108.6172] [INSPIRE].

[21] T. Sjöstrand, *A model for initial state parton showers*, *Phys. Lett. B* **157** (1985) 321 [INSPIRE].

JHEP07(2020)032

[22] M. Ritzmann, D.A. Kosower and P. Skands, *Antenna showers with hadronic initial states*, *Phys. Lett. B* **718** (2013) 1345 [arXiv:1210.6345] [INSPIRE].

[23] S. Frixione and B.R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029 [hep-ph/0204244] [INSPIRE].

[24] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [hep-ph/0409146] [INSPIRE].

[25] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070 [arXiv:0709.2092] [INSPIRE].

[26] S. Catani, F. Krauss, R. Kuhn and B.R. Webber, *QCD matrix elements + parton showers*, *JHEP* **11** (2001) 063 [hep-ph/0109231] [INSPIRE].

[27] L. Lönnblad, *Correcting the color dipole cascade model with fixed order matrix elements*, *JHEP* **05** (2002) 046 [hep-ph/0112284] [INSPIRE].

[28] L. Lönnblad and S. Prestel, *Matching tree-level matrix elements with interleaved showers*, *JHEP* **03** (2012) 019 [arXiv:1109.4829] [INSPIRE].

[29] M.L. Mangano, M. Moretti and R. Pittau, *Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b}$ + n jets as a case study*, *Nucl. Phys. B* **632** (2002) 343 [hep-ph/0108069] [INSPIRE].

[30] M.L. Mangano, M. Moretti, F. Piccinini and M. Treccani, *Matching matrix elements and shower evolution for top-quark production in hadronic collisions*, *JHEP* **01** (2007) 013 [hep-ph/0611129] [INSPIRE].

[31] L. Lönnblad and S. Prestel, *Unitarising matrix element + parton shower merging*, *JHEP* **02** (2013) 094 [arXiv:1211.4827] [INSPIRE].

[32] L. Lönnblad and S. Prestel, *Merging multi-leg NLO matrix elements with parton showers*, *JHEP* **03** (2013) 166 [arXiv:1211.7278] [INSPIRE].

[33] T. Gehrmann, S. Hoche, F. Krauss, M. Schonherr and F. Siegert, *NLO QCD matrix elements + parton showers in $e^+e^- \to$ hadrons*, *JHEP* **01** (2013) 144 [arXiv:1207.5031] [INSPIRE].

[34] S. Hoeche, F. Krauss, M. Schonherr and F. Siegert, *QCD matrix elements + parton showers: the NLO case*, *JHEP* **04** (2013) 027 [arXiv:1207.5030] [INSPIRE].

[35] R. Frederix and S. Frixione, *Merging meets matching in MC@NLO*, *JHEP* **12** (2012) 061 [arXiv:1209.6215] [INSPIRE].

[36] K. Hamilton, P. Nason and G. Zanderighi, *MINLO: multi-scale improved NLO*, *JHEP* **10** (2012) 155 [arXiv:1206.3572] [INSPIRE].

[37] P.F. Monni, P. Nason, E. Re, M. Wiesemann and G. Zanderighi, *MiNNLO$_{\rm PS}$: a new method to match NNLO QCD to parton showers*, *JHEP* **05** (2020) 143 [arXiv:1908.06987] [INSPIRE].

[38] J.J. Lopez-Villarejo and P.Z. Skands, *Efficient matrix-element matching with sector showers*, *JHEP* **11** (2011) 150 [arXiv:1109.3608] [INSPIRE].

[39] N. Fischer and S. Prestel, *Combining states without scale hierarchies with ordered parton showers*, *Eur. Phys. J. C* **77** (2017) 601 [arXiv:1706.06218] [INSPIRE].

[40] S. Höche, S. Prestel and H. Schulz, *Simulation of vector boson plus many jet final states at the high luminosity LHC*, *Phys. Rev. D* **100** (2019) 014024 [arXiv:1905.05120] [INSPIRE].

JHEP07(2020)032

[41] L. Hartgring, E. Laenen and P. Skands, *Antenna showers with one-loop matrix elements*, *JHEP* **10** (2013) 127 [arXiv:1303.4974] [INSPIRE].

[42] H.T. Li and P. Skands, *A framework for second-order parton showers*, *Phys. Lett. B* **771** (2017) 59 [arXiv:1611.00013] [INSPIRE].

[43] D.A. Kosower, *Antenna factorization of gauge theory amplitudes*, *Phys. Rev. D* **57** (1998) 5410 [hep-ph/9710213] [INSPIRE].

[44] D.A. Kosower, *Antenna factorization in strongly ordered limits*, *Phys. Rev. D* **71** (2005) 045016 [hep-ph/0311272] [INSPIRE].

[45] A.J. Larkoski and M.E. Peskin, *Spin-dependent antenna splitting functions*, *Phys. Rev. D* **81** (2010) 054010 [arXiv:0908.2450] [INSPIRE].

[46] A.J. Larkoski and M.E. Peskin, *Antenna splitting functions for massive particles*, *Phys. Rev. D* **84** (2011) 034034 [arXiv:1106.2182] [INSPIRE].

[47] A.J. Larkoski, J.J. Lopez-Villarejo and P. Skands, *Helicity-dependent showers and matching with VINCIA*, *Phys. Rev. D* **87** (2013) 054033 [arXiv:1301.0933] [INSPIRE].

[48] N. Fischer, A. Lifson and P. Skands, *Helicity antenna showers for hadron colliders*, *Eur. Phys. J. C* **77** (2017) 719 [arXiv:1708.01736] [INSPIRE].

[49] N. Fischer, S. Prestel, M. Ritzmann and P. Skands, *Vincia for hadron colliders*, *Eur. Phys. J. C* **76** (2016) 589 [arXiv:1605.06142] [INSPIRE].

[50] H. Brooks and P. Skands, *Coherent showers in decays of colored resonances*, *Phys. Rev. D* **100** (2019) 076006 [arXiv:1907.08980] [INSPIRE].

[51] P. Skands and R. Verheyen, *Multipole photon radiation in the Vincia parton shower*, arXiv:2002.04939 [INSPIRE].

[52] T. Sjöstrand and P.Z. Skands, *Transverse-momentum-ordered showers and interleaved multiple interactions*, *Eur. Phys. J. C* **39** (2005) 129 [hep-ph/0408302] [INSPIRE].

[53] G. Abelof and A. Gehrmann-De Ridder, *Antenna subtraction for the production of heavy particles at hadron colliders*, *JHEP* **04** (2011) 063 [arXiv:1102.2443] [INSPIRE].

[54] F. Krauss and D. Napoletano, *Towards a fully massive five-flavor scheme*, *Phys. Rev. D* **98** (2018) 096002 [arXiv:1712.06832] [INSPIRE].

[55] A. Daleo, T. Gehrmann and D. Maître, *Antenna subtraction with hadronic initial states*, *JHEP* **04** (2007) 016 [hep-ph/0612257] [INSPIRE].

[56] R. Verheyen, *Electroweak effects in antenna parton showers*, Ph.D. thesis, Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen, The Netherlands (2019).

[57] T. Sjöstrand, S. Mrenna and P.Z. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026 [hep-ph/0603175] [INSPIRE].

[58] A. Buckley et al., *Rivet user manual*, *Comput. Phys. Commun.* **184** (2013) 2803 [arXiv:1003.0694] [INSPIRE].

[59] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [arXiv:1405.0301] [INSPIRE].

JHEP07(2020)032

[60] R. Kleiss, W. Stirling and S.D. Ellis, *A new Monte Carlo treatment of multiparticle phase space at high-energies*, *Comput. Phys. Commun.* **40** (1986) 359 [INSPIRE].

[61] B.R. Webber, *Monte Carlo simulation of hard hadronic processes*, *Ann. Rev. Nucl. Part. Sci.* **36** (1986) 253.

[62] T. Plehn, D. Rainwater and P.Z. Skands, *Squark and gluino production with jets*, *Phys. Lett. B* **645** (2007) 217 [hep-ph/0510144] [INSPIRE].

[63] S. Catani, B.R. Webber and G. Marchesini, *QCD coherent branching and semiinclusive processes at large x*, *Nucl. Phys. B* **349** (1991) 635 [INSPIRE].

[64] P. Skands, S. Carrazza and J. Rojo, *Tuning PYTHIA 8.1: the Monash 2013 Tune*, *Eur. Phys. J. C* **74** (2014) 3024 [arXiv:1404.5630] [INSPIRE].

[65] L3 collaboration, *Studies of hadronic event structure in e^+e^- annihilation from 30 GeV to 209 GeV with the L3 detector*, *Phys. Rept.* **399** (2004) 71 [hep-ex/0406049] [INSPIRE].

[66] D0 collaboration, *Precise Study of the Z/γ^* boson transverse momentum distribution in $p\bar{p}$ collisions using a novel technique*, *Phys. Rev. Lett.* **106** (2011) 122001 [arXiv:1010.0262] [INSPIRE].

[67] R. Ellis, D.A. Ross and A.E. Terrano, *The perturbative calculation of jet structure in e^+e^- annihilation*, *Nucl. Phys. B* **178** (1981) 421 [INSPIRE].

[68] J.F. Donoghue, F.E. Low and S.-Y. Pi, *Tensor analysis of hadronic jets in quantum chromodynamics*, *Phys. Rev. D* **20** (1979) 2759 [INSPIRE].

[69] G. Parisi, *Super inclusive cross-sections*, *Phys. Lett. B* **74** (1978) 65 [INSPIRE].

[70] D.P. Barber et al., *Discovery of three jet events and a test of quantum chromodynamics at PETRA energies*, *Phys. Rev. Lett.* **43** (1979) 830 [INSPIRE].

[71] E. Farhi, *A QCD test for jets*, *Phys. Rev. Lett.* **39** (1977) 1587 [INSPIRE].

[72] 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) 57 [INSPIRE].

[73] ATLAS collaboration, *Measurement of dijet azimuthal decorrelations in pp collisions at $\sqrt{s} = 7$ TeV*, *Phys. Rev. Lett.* **106** (2011) 172002 [arXiv:1102.2696] [INSPIRE].

[74] A. Banfi et al., *Optimisation of variables for studying dilepton transverse momentum distributions at hadron colliders*, *Eur. Phys. J. C* **71** (2011) 1600 [arXiv:1009.1580] [INSPIRE].

[75] ATLAS collaboration, *Study of jet shapes in inclusive jet production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, *Phys. Rev. D* **83** (2011) 052003 [arXiv:1101.0070] [INSPIRE].

[76] M. Wobisch and K. Rabbertz, *Dijet azimuthal decorrelations for $\Delta\varphi_{\text{dijet}} < 2\pi/3$ in perturbative QCD*, *JHEP* **12** (2015) 024 [arXiv:1505.05030] [INSPIRE].

[77] S. Ferrario Ravasio, T.s. Jeˇ zo, P. Nason and C. Oleari, *A theoretical study of top-mass measurements at the LHC using NLO+PS generators of increasing accuracy*, *Eur. Phys. J. C* **78** (2018) 458 [*Addendum ibid.* **79** (2019) 859] [arXiv:1906.09166] [INSPIRE].

[78] S. Höche and S. Prestel, *Triple collinear emissions in parton showers*, *Phys. Rev. D* **96** (2017) 074017 [arXiv:1705.00742] [INSPIRE].

JHEP07(2020)032

[79] S. Höche, F. Krauss and S. Prestel, *Implementing NLO DGLAP evolution in parton showers*, *JHEP* **10** (2017) 093 [arXiv:1705.00982] [INSPIRE].

[80] Z. Nagy and D.E. Soper, *What is a parton shower?*, *Phys. Rev. D* **98** (2018) 014034 [arXiv:1705.08093] [INSPIRE].

[81] F. Dulat, S. Höche and S. Prestel, *Leading-color fully differential two-loop soft corrections to QCD dipole showers*, *Phys. Rev. D* **98** (2018) 074013 [arXiv:1805.03757] [INSPIRE].

[82] M. Dasgupta et al., *Logarithmic accuracy of parton showers: a fixed-order study*, *JHEP* **09** (2018) 033 [*Erratum ibid.* **03** (2020) 083] [arXiv:1805.09327] [INSPIRE].

[83] M. Dasgupta et al., *Parton showers beyond leading logarithmic accuracy*, arXiv:2002.11114 [INSPIRE].

[84] R. Kleiss and R. Verheyen, *Electroweak radiation in antenna parton showers*, arXiv:2002.09248 [INSPIRE].

JHEP07(2020)032

5

Efficient Tree-Level Merging with Sector Showers

Parton showers give an exclusive description of collider events normalised to the Born-level LO or NLO (if the shower has been matched) cross section. The description of additional radiation by the shower is, however, only accurate in the soft and collinear limits, where the shower resums large logarithms, *cf.* section 3.2. Away from these, in regions where hard, well-separated jets dominate, fixed-order calculations yield accurate predictions, *cf.* section 3.1.

A well-established technique to correct parton showers to tree-level matrix element in hard phase-space regions is given by the CKKW-L merging scheme [346–349], *cf.* section 3.3.3. In the CKKW-L scheme, a (somewhat arbitrary) merging scale t_{MS} is introduced, which separates the parton-shower region ($t \leq t_{\mathrm{MS}}$) from the matrix-element region ($t > t_{\mathrm{MS}}$). Both calculations are then only allowed to populate their designated regions, which is achieved via vetoing branchings in the former and vetoing configurations in the latter. In addition, the Sudakov (resummation) suppression inherent to parton showers has to be accounted for in the fixed-order matrix elements to render the initially inclusive calculation exclusive. To this end, a parton shower history has to be constructed for each matrix-element configuration, so that trial showers between the nodes can be used to calculate the no-branching probabilities in exactly the same way as the shower at hand would.

As alluded to in chapter 4, the number of shower histories grows factorially with the number of particles. As a result, the CKKW-L scheme inherits a rather dramatic scaling with the number of jets, which directly translates into the event-generation time and the memory footprint of the method. The latter arises because all possible shower histories have to be constructed before the most-likely (according to the shower branching probabilities) can be picked, which implies that a large number of histories needs to be stored. Excessive memory allocation is the most restricting bottleneck of merging at high multiplicities. Despite these restrictions, merged predictions for Drell-Yan processes with up to nine additional jets were obtained in [145]. However, for configurations with more than six jets, the shower history was there only approximated via a deterministic jet-clustering algorithm.

Sector showers offer a resort to these problems, as the shower evolution is constrained by a deterministic shower-based jet-clustering algorithm, which renders sector showers maximally bijective. This means, that states with maximally one quark-antiquark pair

can be uniquely inverted, whereas states with $n_{q\bar{q}}$ quark pairs have a maximum of $n_{q\bar{q}}!$ competing histories. Consequently, merging with sector showers is exceptionally simple: any given (shower-history-based) merging scheme can be used and amended by a dedicated algorithm for the construction of the sector-shower history.

In the publication included in section 5.1, a dedicated merging scheme for sector showers is discussed. The new technique combines the CKKW-L algorithm with the maximally-bijective nature of sector showers, resulting in an efficient tree-level multi-jet merging method which nevertheless exactly accounts for the shower no-branching probabilities and the phase space it populates. The new algorithm is implemented for VINCIA's sector showers in PYTHIA 8.3, largely independent of PYTHIA's default merging implementation. Using the HDF5 event files of [423–425], it is demonstrated that the complexity of constructing the shower history develops an effective linear scaling with the number of additional jets and that both the event generation time and the memory allocation remain approximately constant when including higher jet multiplicities.

5.1 Published Material

Efficient multi-jet merging with the Vincia sector shower

Helen Brooks and **Christian T Preuss**

Published in Computer Physics Communications 264 (2021) 107985
DOI: 10.1016/j.cpc.2021.107985
e-Print: arXiv:2008.09468 [hep-ph]

Publication begins overleaf.

Computer Physics Communications 264 (2021) 107985

Contents lists available at ScienceDirect

Computer Physics Communications

www.elsevier.com/locate/cpc

Efficient multi-jet merging with the VINCIA sector shower ☆

Helen Brooks [1], Christian T. Preuss *

School of Physics and Astronomy, Monash University, Wellington Road, Clayton, VIC-3800, Australia

ARTICLE INFO

Article history:
Received 18 September 2020
Received in revised form 8 March 2021
Accepted 2 April 2021
Available online xxxx

Keywords:
QCD jets
Parton showers
Multi-jet merging

ABSTRACT

We here present an extension of the CKKW-L multi-jet merging technique to so-called sector showers as implemented in the VINCIA antenna shower. The bijective nature of sector showers allows for efficient multi-jet merging at high multiplicities, as any given configuration possesses only a single "history", while retaining the accuracy of the CKKW-L technique. Our method reduces the factorial scaling of the number of parton shower histories to a constant of a single history per colour-ordered final state. We show that the complexity of constructing shower histories is reduced to an effective linear scaling with the number of final-state particles. Moreover, we demonstrate that the overall event generation time and the memory footprint of our implementation remain approximately constant when including additional jets. We compare both to the conventional CKKW-L implementation in PYTHIA and gain a first estimate of renormalisation scale uncertainties at high merged multiplicities. As a proof of concept, we show parton-level predictions for vector boson production in proton-proton collisions with up to nine hard jets using the new implementation. Despite its much simpler nature, we dub the new technique MESS, in analogy to the conventional MEPS nomenclature.

Contents

1. Introduction

While fixed-order calculations accurately describe observables in regions of phase space where hard, well-separated jets dominate, they are insufficient in the resummation region, where additional particles are emitted at low energies or angles. In these collinear- and soft-enhanced phase space regions, parton shower

☆ The review of this paper was arranged by Prof. Z. Was.

* Corresponding author.
E-mail address: christian.preuss@monash.edu (C.T. Preuss).
[1] Now at CCFE, Oxfordshire, UK.

https://doi.org/10.1016/j.cpc.2021.107985
0010-4655/

Monte Carlo generators provide a reliable and versatile tool to resum the leading logarithms (LL) arising from QCD matrix elements to all orders in the perturbative expansion in the strong coupling. In order to achieve an accurate description over all of phase space, these two approaches need to be combined, by techniques known as matching or merging. With experimental analyses becoming available for high jet multiplicities, cf. e.g. [1,2], and in the advent of the high-luminosity LHC, the demand for calculations with both an accurate description of many hard jets as well as QCD bremsstrahlung is ever increasing.

To date, a vast amount of matching and merging schemes has been developed, with matching to leading-order (LO) [3–6] or next-to-leading-order (NLO) [7–9] matrix elements on the one hand and merging with LO [10–18] and NLO [19–26] calculations on the other. First steps towards the inclusion of NNLO calculations have also been taken [27–33]. Together with the automation of tree-level matrix-element generation [34–41] the path has been paved for tree-level matching and merging at high multiplicities. However, the computational overhead of such calculations grows at least factorially with the number of particles on both the fixed-order and the resummation side, quickly rendering such computations intractable.

Although the most restrictive bottlenecks in merged calculations arise in the context of generating high-multiplicity matrix elements [42], especially the phase-space integration, these parton-level samples can be generated "once and for all", meaning they can be re-utilised for many different particle-level analyses, given the generation is sufficiently factorised. A novel framework for factorised fixed-order and parton-shower calculations in the high-multiplicity regime has been presented in [42]. The parton-level events of vector boson production with up to 9 additional jets generated there have been made publicly available [43–45].

Leaving the difficulties and pitfalls of high-multiplicity matrix-element generation aside, the main bottleneck in merged calculations then arises from the fact that merging schemes usually rely on the construction of parton shower histories, i.e., the sequences of states the parton shower would have produced to arrive at a given configuration. The purpose of constructing parton shower histories is to obtain Sudakov factors to reweight *inclusive* event samples to make them *exclusive*, so that double-counting of emissions is avoided. Hence, this procedure has to be undertaken for every parton-shower merged calculation and therefore particle-level event generation run.

In conventional dipole- or DGLAP-based parton showers, the number of histories grows at least factorially with the number of final-state particles. Naturally, one can resort to a deterministic (sometimes referred to as "winner-takes-all") scheme, where a simple jet-clustering algorithm is employed to construct the shower history. Such a scheme can, however, in principle lead to under- or over-counted phase space regions and consequently may not correctly reflect the Sudakov factors generated by the shower. Constructing and weighting all possible histories therefore becomes highly time- and resource-intensive for large final-state multiplicities. Moreover, the memory required to store all possible histories until the most probable is picked may exceed the available memory, cf. [42]. In that study, the construction of all possible histories was therefore limited to up to 6 additional jets, after which a deterministic ("WTA") scheme was employed. Although the effect was not found to be large, in a precision calculation it is desirable that the Sudakov factors exactly match the ones in the shower evolution.

We here present a new implementation to combine tree-level matrix elements with so-called sector showers [46,47], based on the CKKW-L merging prescription [12,13,48,16]. In the sector shower framework, only a single splitting kernel contributes at any point in phase space, making the shower operator effectively bijective, i.e., uniquely invertible, while retaining its leading-logarithmic accuracy. Hence, for any given multi-parton configuration, there exists only a single path to every previous shower state and the factorially growing history tree is replaced by a single, linear history branch. The method presented here, which, in analogy to the MEPS nomenclature, we dub MESS as a shorthand for matrix elements + sector shower, alleviates the scaling of the memory footprint as well as computation time on the parton shower side of multi-jet merging. The MESS presented here was made publicly available in the PYTHIA 8.304 release.

This paper is structured as follows. We review the CKKW-L merging scheme for the VINCIA sector shower with a particular focus on the construction of shower histories in Section 2 and validate the new implementation in Section 3. Our central results – the run time scaling and memory footprint of our implementation – are presented in Section 4 before concluding in Section 5.

2. CKKW-L merging with sector showers

Generally, we shall be performing tree-level CKKW-L merging [13], following the prescription in [16] with slight modifications to adapt it to sector showers. Although this method is in principle more generally applicable, we will limit our discussion to the VINCIA antenna shower, for which the extension to sector showers was discussed in detail in [47]. We will briefly comment on the sectorisation of other shower approaches in Section 2.2. For a review of the CKKW and the CKKW-L approach, we refer to [49,16]; here it shall suffice to present a brief review only.

2.1. CKKW-L merging in a nutshell

To safely combine multiple event samples, including the effect of shower simulations, without over-counting emissions, the initially inclusive events have to be made exclusive. In order to do so, first a "merging scale", t_{MS}, has to be defined. It is used to separate the fixed-order and resummation regions, so that each shower-generated jet falls below the merging scale and each matrix-element generated jet falls above it. Starting from a given Born$+n$-jet event that passes this constraint, a typical merging algorithm can then be separated into three steps:

1. Construct the most likely "shower history" consisting of sequential clustering "nodes"
2. Reweight the event with Sudakov factors to account for unresolved radiation
3. Reweight the event with α_s factors, evaluated at appropriate "node scales"

The idea of the CKKW-L merging prescription as presented in [13,16] is to generate the Sudakov factors in the second step dynamically and in the same way as the parton shower at hand would have done while reaching the given Born $+ n$ configuration, as described below.

For a configuration with n additional jets with respect to the Born configuration, which has been generated according to a tree-level matrix element with a regularisation cutoff k_{cut}, the most probable shower history is reconstructed in the first step. Denoting the hard event by $\mathcal{H}_{\mathrm{Born}+n}$ and shower states by $\mathcal{S}_{\mathrm{Born}+i}$, this generates a sequence of nodes

$$\{\mathcal{S}_{\mathrm{Born}}, \mathcal{S}_{\mathrm{Born}+1}, \ldots, \mathcal{S}_{\mathrm{Born}+n-1}, \mathcal{H}_{\mathrm{Born}+n}\}\,, \tag{1}$$

with a corresponding sequence of node scales, $\{\rho_0, \rho_1, \ldots, \rho_{n-1}, \rho_n\}$, typically given by the shower evolution variable.

Subsequently, Sudakov form factors are generated by trial showers between history nodes, during which an event is vetoed if a branching between two nodes is produced. This generates

no-branching probabilities $\Pi_{\mathcal{S}_{\text{Born}+i}}(\rho_i, \rho_{i+1})$ in the same way the shower had if it would have been started off the reconstructed nodes. As the no-branching probability $\Pi_{\mathcal{S}_i}$ generated by the shower generally differs from Sudakov factors by PDF ratios, the events are additionally weighted by

$$w_i^{\text{PDF}} = \frac{f_i(x_i, \rho_i)}{f_i(x_i, \rho_{i+1})}\,. \tag{2}$$

To account for the running of the strong coupling and other higher-order corrections included in the shower evolution, events are weighted with ratios

$$w_i^{\alpha_s} = \frac{\alpha_{s,\text{PS}}(\rho_i)}{\alpha_{s,\text{ME}}} \tag{3}$$

for each intermediate node, where $\alpha_{s,\text{PS}}$ and $\alpha_{s,\text{ME}}$ reflect the scale and scheme choice of the shower and fixed-order calculation, respectively.

In the last trial shower step, the treatment differs between intermediate and highest-multiplicity nodes. The event is vetoed if the trial shower off the hard $\mathcal{H}_{\text{Born}+n}$ configuration generates an emission above the merging scale, $t_n(\mathcal{H}_{\text{Born}+n}) > t_{\text{MS}}$ and n is below the maximal number of additional jets N. Hard emissions off configurations with the highest jet multiplicity $n \equiv N$ are, however, retained. Here, the notation $t_n(\mathcal{S})$ denotes the evaluation of the state $\mathcal{S}$ with respect to the same metric as used for the merging scale t_{MS}. This can be a simple jet-$p_\perp$ cut, the shower evolution variable, or more complicated definitions including the use of jet clustering algorithms.

A hard Born + n parton configuration is therefore accepted if and only if the trial showers did not generate additional hard emissions. Accepted events are thus weighted by

$$\begin{aligned} w_{\text{CKKW-L}} = \frac{f_n(x_n, \rho_n)}{f_n(x_n, \mu_F^2)} \prod_{i=0}^{n-1} \frac{\alpha_{s,\text{PS}}(\rho_{i+1})}{\alpha_{s,\text{ME}}} \frac{f_i(x_i, \rho_i)}{f_i(x_i, \rho_{i+1})} \\ \times \Pi_{\mathcal{S}_{\text{Born}+i}}(\rho_i, \rho_{i+1})\,, \end{aligned} \tag{4}$$

where μ_F^2 denotes the factorisation scale of the hard process.

2.2. Sectorised shower evolution

The VINCIA antenna showers are evolved in a generalised ARIADNE $p_\perp$,

$$\begin{aligned} p_\perp^2 &= \frac{\bar{q}_{ij}^2 \bar{q}_{jk}^2}{s_{\max}}\,, \\ \bar{q}_{ij}^2 &= \pm[(p_i \pm p_j)^2 - p_I^2] \\ &= \begin{cases} s_{ij} + m_i^2 + m_j^2 - m_I^2 & i \text{ is final} \\ s_{ij} - m_i^2 - m_j^2 + m_I^2 & i \text{ is initial} \end{cases}\,, \end{aligned} \tag{5}$$

where capital indices denote pre-branching partons and $s_{ij} = 2p_i \cdot p_j$ with $s_{\max}$ the maximal invariant of the current antenna,

$$s_{\max} = \begin{cases} s_{IK} & \text{final-final} \\ s_{ij} + s_{ik} & \text{initial-final} \\ s_{ik} & \text{initial-initial} \end{cases}\,. \tag{6}$$

In the sector shower formalism, only a single antenna contributes at each point in phase space. In order to nevertheless capture the correct leading-logarithmic behaviour, a single antenna function incorporates both the full soft and the full collinear singularity. The exact form of sector antenna functions is, however, ambiguous and only limited by the constraint that the correct single-unresolved limits are entirely contained within a single function. We refer to [47] for a full set of helicity- and mass-dependent sector antenna functions, which, in its colour factor- and coupling-stripped variant, we denote by $\bar{a}^{\text{sct}}$ here. The shower operator is made bijective by rejecting any branching that does not correspond to the most singular configuration in the tentative post-branching state, cf. Fig. 1, defined in terms of the sector resolution variable

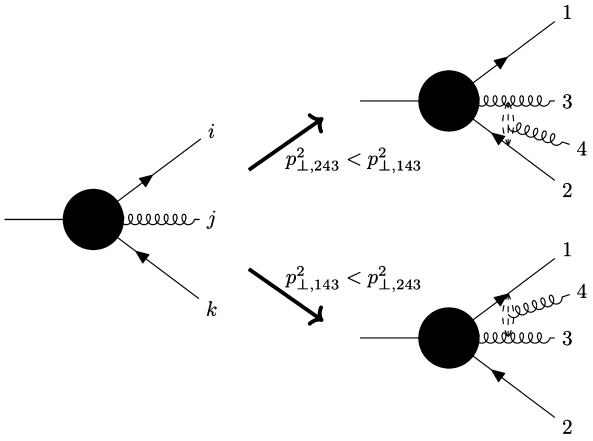


Fig. 1. Illustration of the sector-shower evolution off a colour-ordered $Z \to qg\bar{q}$ configuration. The emission in the i-j antenna is accepted if and only if its post-branching $p_\perp$ is the smallest in the tentative post-branching state.

$$Q_{\text{res}_j}^2 = \begin{cases} p_\perp^2 & \text{if } j \text{ is a gluon} \\ \bar{q}_{ij}^2 \sqrt{\dfrac{\bar{q}_{jk}^2}{s_{\max}}} & \text{if } j \text{ is a(n) (anti)quark} \end{cases}\,. \tag{7}$$

Here, the asymmetric choice for gluon splittings accounts for the fact that the $g_I X_K \mapsto q_i \bar{q}_j X_k$ branching is not singular in the j-k collinear limit, cf. [46]. Thus, the sector shower produces no-branching probabilities of the form

$$\begin{aligned} &\Pi_n(p_{\perp,n}^2, p_{\perp,n+1}^2) \\ &= \exp\left(-4\pi \sum_{j \in \{n \mapsto n+1\}} \int_{p_{\perp,n+1}^2}^{p_{\perp,n}^2} \frac{f_i(x_i, p_\perp^2) f_k(x_k, p_\perp^2)}{f_I(x_I, p_\perp^2) f_K(x_K, p_\perp^2)} \alpha_s(p_\perp^2) \right. \\ &\qquad \left. \times C_{j/IK} \bar{a}_{j/IK}(p_\perp^2, \zeta)\, \Theta_j^{\text{sct}}(p_\perp^2, \zeta)\, \mathrm{d}\Phi_j^{\text{ant}} \right), \end{aligned} \tag{8}$$

where the Heaviside function Θ^{sct} enforces the constraint that only a single antenna radiates per phase-space point. In general, it depends non-trivially on the post-branching kinematics.

Consequently, any given configuration produced by the sector shower can be uniquely inverted by iteratively minimising Eq. (7), effectively yielding a $p_\perp^2$-based jet-clustering algorithm, which, however, still *exactly* represents the (leading-colour) shower history.

Before moving on to the treatment of shower histories, it is worthwhile to discuss how other shower models may be extended to sector showers. Compared to conventional shower algorithms, sector showers differ mainly in the choice of branching kernels and their phase-space coverage. These are constructed based on the following two requirements [50,51]:

1. a single branching kernel contains all single-unresolved limits of the respective colour dipole

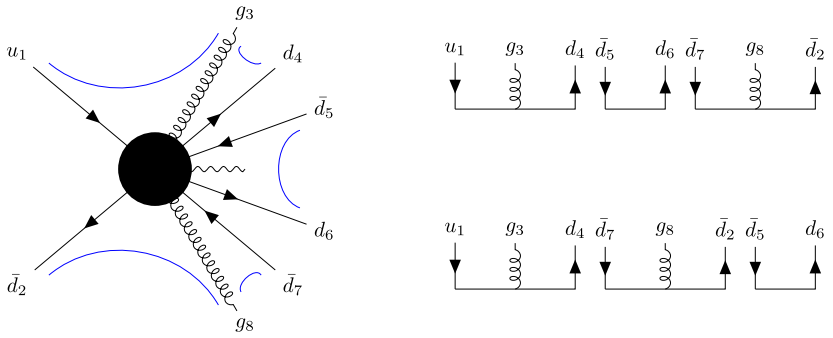


Fig. 2. Illustration of colour chains used for history construction. For the colour-ordered configuration (*left*), two different permutations of the same three colour chains contribute (*right*). Both constitute one sector shower history.

2. the branching phase space is decomposed into non-overlapping sectors, where only a single branching kernel contributes per sector and the set of sectors provides a decomposition of unity

Other aspects, which would typically be left unchanged when sectorising a shower model, but which will of course affect the resummation of leading logarithms include a judicious choice of the ordering variable [52–56], correct assignment of colour factors [57–62], and the treatment of branching recoils [56,62]. The choices made in the VINCIA sector shower are discussed in detail in [47].

As a second specific example, consider a shower based on Catani-Seymour dipoles [63,64]. As opposed to (global) antenna showers, in which the gluon-gluon collinear limits are partitioned between two neighbouring antennae, in Catani-Seymour dipoles it is the soft limit which is partitioned between the two dipole legs. Sector splitting kernels can therefore be constructed by combining two Catani-Seymour dipoles, e.g. for gluon emission from a quark-gluon dipole:

$$V^{\text{sct}}_{q_i g_j, g_k}(z_i, z_k, y_{ij,k}) = \frac{V_{q_i g_j, k}(z_i, y_{ij,k})}{s_{ij}} + \frac{V_{g_k g_j, i}(z_k, y_{jk,i})}{s_{jk}}\,, \quad (9)$$

where $V_{ij,k}$ denotes the colour-stripped versions of the spin-averaged splitting functions $\langle \boldsymbol{V}_{ij,k} \rangle$ in [63] and

$$z_i = \frac{s_{ik}}{s_{ik} + s_{jk}}\,, \quad y_{ij,k} = \frac{s_{ij}}{s_{ij} + s_{jk} + s_{ik}}\,. \quad (10)$$

Splitting functions constructed this way fulfil the first requirement above, as they approach the soft eikonal and the correct DGLAP kernels in the respective soft and collinear limits. Supplemented with a sensible measure of singularity, for instance taken to be $k^2_\perp = z_i(1 - z_i)s_{ij}$, they can be restricted to the appropriate phase space sectors in the same way as in Eq. (8),

$$1 = \sum_j \theta(k^2_{\perp,\min} - k^2_{\perp,j})\,, \quad (11)$$

therefore fulfilling the second requirement above.

2.3. Shower histories

The advantage of sector showers is that, at least for gluon emissions at leading colour, there is just a single history, because only one antenna is active at each point in phase space. Thus, the history for gluon emissions may be constructed deterministically by minimising the resolution criterion Eq. (7).

For quark pairs, however, the situation is less clear-cut. While for gluons, we can use their colour-connected neighbours to determine their possible parents (and the sector criterion determines the order in which they were emitted), for quark-antiquark pairs, there is no colour information to determine which pairs should be clustered. Therefore, in principle we must consider all clusterings of opposite-sign same-flavour pairs.[2] We can do this by taking all possible orderings of the colour-ordered chains of gluons, each of which starts on a quark and ends on an antiquark, such that quarks which are juxtaposed are of the same flavour, and clustering all pairs of quark-antiquark pairs results in a viable Born-level topology, cf. Fig. 2. Nevertheless, this results in a significant gain in efficiency, since the number of orderings only grows as $\mathcal{O}\left(\prod_i (n_{q_i}!)\right)$, where n_{q_i} is the number of pairs of quark flavour i, which is at worst $n/2$, but typically much smaller. We emphasise that once a colour ordering of the quark pairs is picked, the shower history is again deterministic.

As we are only trying to capture the leading singular behaviour of the matrix element, we calculate

$$|\mathcal{M}_{\text{Born}}|^2 \prod_{i=1}^{n} \bar{a}_i^{\text{sct}}(\{p\}_i) \propto |\mathcal{M}_{\text{Born}+n}|^2\,, \quad (12)$$

for each viable colour ordering and then maximise over this quantity. Therefore we only need to save maximally two histories concurrently: the current one plus the "best-so-far", i.e., that which maximises Eq. (12).

In summary, the sector shower history is constructed as follows:

1. Find all colour-connected chains of gluons.
2. Find all possible orderings of colour chains compatible with the Born-level process.
3. For each available permutation:
   - Sequentially perform the clustering which corresponds to the minimal value of the resolution criterion Q^2_{res}, cf. Eq. (7). For each state $\mathcal{S}_n$:
     - For all gluons and internal quark pairs, calculate the sector resolution variable. Note that for quarks, there is an ambiguity in the recoiler, so there are two antennae per quark pair.
     - Cluster the partons which correspond to the minimal value of the sector resolution variable Eq. (7).
     - Reconstruct the $(n-1)$-parton kinematics using the exact inverse kinematics map, cf. [47].
     - Store mother/daughter information.
     - Update the colour chain information.
     - Calculate the evolution variable Eq. (5) for the branching (in general not the same as the sector resolution variable).

[2] Assuming that there are no flavour-changing (i.e. electroweak) emissions from the shower.

  - Calculate the sector antenna function from the invariants of the pre-clustering partons.
- Retain the history for the current permutation only if it corresponds to the maximal value of Eq. (12) so far.

Below, we address some subtleties connected to the construction of sector shower histories.

Commensurate Resolution Scales

We construct the sector shower history by minimising the sector resolution variable Eq. (7). In principle, however, two (or more) different clusterings may have very similar sector resolutions and it may seem unreasonable to choose one over the other. In those cases, one could consider to pick one of the commensurate-resolution clusterings randomly. However, given that this would destroy the merit of the sector shower being uniquely invertible, as the same would have to be done in the shower evolution, we refrain from this procedure and always pick the one with the (slightly) smaller scale. Given that this precisely mimicks the sector shower behaviour, this is a well-motivated choice. Nevertheless, the effect of randomising the choice of such clusterings is an interesting subject for a later study and could be used for uncertainty estimates.

Unordered Histories

Although the history is constructed based on minimising the sector resolution, Eq. (7), there is no guarantee that this produces a clustering sequence ordered in either the resolution or the evolution variable, Eq. (5). As the sector shower is based purely on $2 \mapsto 3$ branchings, it will never populate regions of phase space with branchings unordered in the evolution variable. Hence, no Sudakov factors must be included for unordered (sub-)sequences. This is similar to the treatment in [16]; we note also that an unordered history is only selected if no ordered one exists.

Incomplete Histories

Occasionally it may occur that it is impossible to perform any (further) parton shower clusterings. Physically, these topologies correspond to states that cannot be reached by the parton shower from any lower multiplicity state. Therefore, there is no danger of double-counting with the lower-multiplicity states, and these states are treated as coming from separate Born configurations. In the event that there are multiple colour histories, we must calculate a modification to Eq. (12) as our criterion to maximise, namely:

$$|\mathcal{M}_{\text{Born+m}}|^2 \prod_{i=m}^{n} \bar{a}_i^{\text{sct}}(\{p\}_i) \propto |\mathcal{M}_{\text{Born}+n}|^2, \tag{13}$$

where m is the number of additional emissions relative to the Born in the maximally clustered node of the incomplete history. This still allows to select the most singular path, since in effect this compares $|\mathcal{M}_{\text{Born}+n}|^2$ with $|\mathcal{M}_{\text{Born}+m}|^2 \prod_{i=m}^{n} \bar{a}_i^{\text{sct}}(\{p\}_i)$; since the latter captures the singularity structure of the former, it is a fair comparison. We follow the procedure of [16] and accept an incomplete history only if no colour permutation with a complete one exists.

Interleaved Multi-Parton Interactions

In the context of interleaved showers for hadronic initial states, it is possible that the trial shower may generate a multi-parton interaction (MPI) "emission" from an intermediate clustered state in the history. Since such topologies are not reachable by the matrix-element and it would not be physical to limit the scale of MPI to below the merging scale, such "new" topologies are taken to replace the original event, and showering continues from the scale at which the MPI was generated. This is precisely the same treatment as [16].

Scale definitions

In the sector shower merging algorithm, up to four different scale definitions may be present:

1. the shower evolution variable $p_\perp$
2. the sector resolution variable Q_{res}
3. the merging scale t_{MS}
4. the matrix element cutoff k_{cut}

The sector resolution variable Q_{res} is only used to construct the shower history and does not play a role in the merging algorithm beyond that. If the other three scales are not chosen to coincide, care has to be taken to neither over- nor undercount phase space regions.

To ensure a smooth transition between the shower evolution variable and the merging scale, we reject hard configurations if any of the intermediate states violates the merging scale cut, i.e. if $t_i(\mathcal{S}_{\text{Born}+i}) < t_{\text{MS}}$. This is different to the implementation in PYTHIA, where intermediate nodes are not required to be above the merging scale, as multiple shower paths contribute to the same phase space point. We refrain from this treatment, as in our case, given the same hard configuration multiple times, the sector shower history will always be the same. This, however, is not sufficient to ensure that the hard phase space is saturated when the matrix element cutoff is chosen with respect to a different scale definition than the merging scale. This will only be the case when the available phase space with the merging scale cut is a subset of the phase space with the matrix element cut, $\Phi_{>t_{\text{MS}}} \subset \Phi_{>k_{\text{cut}}}$.

2.4. The full algorithm

For completeness, we here summarise the full CKKW-L merging algorithm for sector showers, closely following [16]:

(1) Pick a hard event $\mathcal{H}_{\text{Born}+n}$ containing n additional partons relative to the Born topology:
- If the hard configuration does not satisfy the merging scale cut, i.e., $t_n(\mathcal{H}_{\text{Born}+n}) < t_{\text{MS}}$, veto the event and start from (1).
- For each viable colour-ordering, reconstruct the (deterministic) sequence of shower states $\mathcal{S}_{\text{Born}+i}$,

$$\{\mathcal{S}_{\text{Born}}, \mathcal{S}_{\text{Born}+1}, \ldots, \mathcal{S}_{\text{Born}+n-1}, \mathcal{H}_{\text{Born}+n}\} \tag{14}$$

with a corresponding sequence of evolution variables

$$\left\{p^2_{\perp,0}, p^2_{\perp,1}, \ldots, p^2_{\perp,n-1}, p^2_{\perp,n}\right\}. \tag{15}$$

Here, $p^2_{\perp,i}$ denotes the sector shower evolution scale of the branching to produce each state and $p^2_{\perp,0}$ is the kinematic limit of the Born process, i.e., $p^2_{\perp,1}$ is the scale to produce the first branching, $p^2_{\perp,2}$ is the scale to produce the second branching and so on.
- If any state does not satisfy the merging scale cut, i.e., if $t_i(\mathcal{S}_{\text{Born}+i}) < t_{\text{MS}}$ for any $0 \leq i < n$, veto the event and start from (1).

(2) For each pair of scales $(p^2_{\perp,i}, p^2_{\perp,i+1})$, where $m \leq i < n$ ($m = 0$ for complete histories):
- If the pair is unordered, $p^2_{\perp,i+1} > p^2_{\perp,i}$, move to the next pair.
- Else, perform trial shower between the two scales:

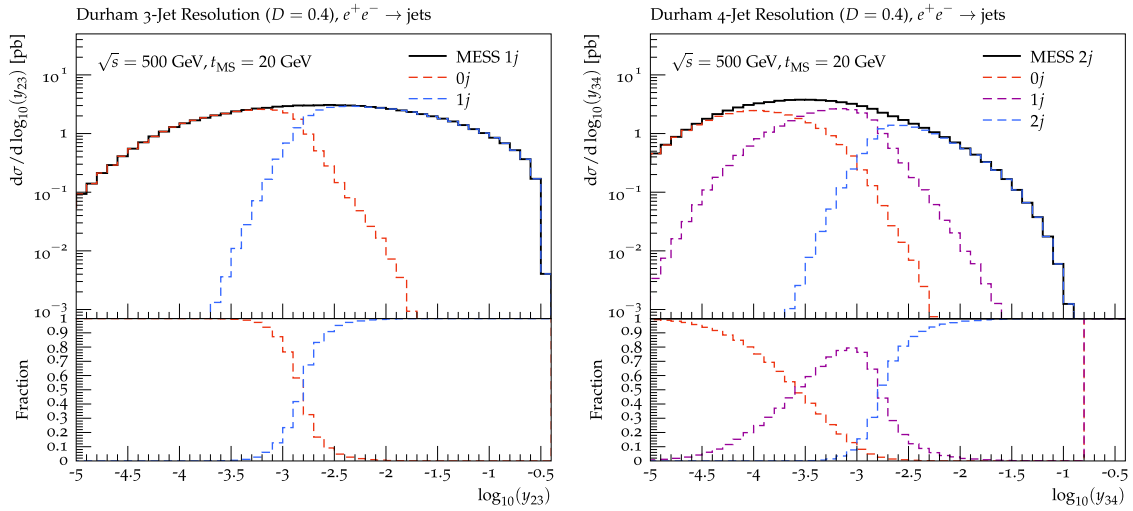


Fig. 3. Contributions of the individual hard-event samples in Durham 3-jet (*left*) and 4-jet (*right*) resolution scales in $e^+e^- \to$ jets at $\sqrt{s} = 500$ GeV. (For interpretation of the colours, the reader is referred to the online version of this article.)

- If the generated state $\mathcal{S}_{\text{Born}+j}$ has a MPI, accept event and move to step (3).
- Else, if $\mathcal{S}_{\text{Born}+j}$ has $p^2_{\perp,j} > p^2_{\perp,i+1}$, veto the event and start from (1).
- Else, calculate the weights

$$w_i^{\alpha_s} = \frac{\alpha_{s,\text{PS}}(p^2_{\perp,i+1})}{\alpha_{s,\text{ME}}},$$
$$w_i^{\text{PDF}} = \frac{f_i^A(x_i^A, p^2_{\perp,i})}{f_i^A(x_i^A, p^2_{\perp,i+1})} \frac{f_i^B(x_i^B, p^2_{\perp,i})}{f_i^B(x_i^B, p^2_{\perp,i+1})} \tag{16}$$

(3) If the event was not vetoed:

- Multiply the event weight by

$$w_{\text{CKKW-L}} = \frac{f_n^A(x_n^A, p^2_{\perp,n})}{f_n^A(x_n^A, \mu_F^2)} \frac{f_n^B(x_n^B, p^2_{\perp,n})}{f_n^B(x_n^B, \mu_F^2)} \prod_{i=0}^{n-1} w_i^{\alpha_s} w_i^{\text{PDF}}. \tag{17}$$

  This is a probabilistic way to generate the factor

$$\prod_{i=0}^{n-1} w_i^{\alpha_s} \Delta(p^2_{\perp,i}, p^2_{\perp,i+1}).$$

- Start the regular parton shower from the state $\mathcal{H}_{\text{Born}+n}$ at scale $p^2_{\perp,n}$. If $n+1 \leq N$ veto the event and start from (1) if $t_{n+1}(\mathcal{S}_{(\text{Born}+n)+1}) > t_{\text{MS}}$.

(4) Start over from (1).

The algorithm outlined above has been implemented for the VINCIA parton shower within the PYTHIA 8.3 event generator. Notwithstanding the use of some existing data structures in PYTHIA, our implementation is largely independent from that of the original CKKW-L merging algorithm.

3. Validation

We validate our implementation in electron-positron annihilation processes and vector boson production in proton-proton collisions. Event samples are generated with the MADGRAPH5_aMC@NLO event generator [41] using the NNPDF23_lo_as_0130_qed PDF set with fixed renormalisation and factorisation scale corresponding to the mass of the Z, $\mu_R = \mu_F = M_Z$. Jets are defined using the $k_\perp$ jet clustering algorithm with a radius parameter of $D = 0.4$ and matrix elements are regularised by a $k_\perp$ cut. To not obscure the effect of the merging, we consider parton-level results only and do not include MPIs. The merging scale is defined by the matrix element $k_\perp$ cut in all cases and we use the default VINCIA tune of the perturbative parameters with $\alpha_s^{\overline{\text{MS}}}(M_Z) = 0.118$, summarised in Appendix A. For electron-positron annihilation we choose a default merging scale of $t_{\text{MS}} = 20$ GeV, while for vector boson production, we choose a lower default merging scale of $t_{\text{MS}} = 10$ GeV, as the sector shower is currently not corrected to matrix elements. Analyses are performed using RIVET [65,66].

Fig. 3 shows the contribution of hard 3-jet and 4-jet events in Durham jet resolution scales in $e^+e^- \to$ jets at a centre-of-mass energy of $\sqrt{s} = 500$ GeV. The effect of varying the merging scale by a factor of two is shown in Fig. 4 for the Durham 3-jet resolution and the C parameter (for a definition see e.g. [67]).

In Fig. 5, the influence of merging the sector-shower predictions with up to two hard matrix elements on $k_\perp$ 1- and 2-jet resolution scales in Drell-Yan processes in the electron channel, $pp \to Z +$jets at $\sqrt{s} = 14$ TeV is studied. The effect of varying the merging scale by a factor of two is presented in Fig. 6.

The individual contributions of the Born, 1-, and 2-jet event samples to $k_\perp$ 1- and 2-jet resolution scales in W boson production in proton-proton collisions at $\sqrt{s} = 14$ TeV are shown in Figs. 7 and 9. Figs. 8 and 10 shows the influence of different merging scale choices on the merged predictions of the $k_\perp$ 1-jet splitting scale and the W boson transverse momentum spectrum.

The jet-resolution scale distributions clearly show the expected behaviour that showers off Born configurations dominate the low-energy region on the left-hand side, while showers off higher-multiplicity states dominate in the hard region towards the right-hand side of the plots. The transition region smoothly interpolates between these two phase-space regions. A variation of the merging scale by factors of two has negligible effects on the distributions in e^+e^- annihilation. For vector boson production in proton-proton collisions, the effect of choosing a higher merging scale results in a more pronounced effect on the distributions. Given that the underlying sector shower starts at the factorisation scale and is not corrected to matrix elements in a process that is subject to sizeable corrections from 1- and 2-jet matrix elements, we deem this a reasonable effect. This argument is supported by the fact that choosing a smaller merging scale has a far less-pronounced effect.

We have verified that the above statements remain true for a significantly larger set of observables than presented here.

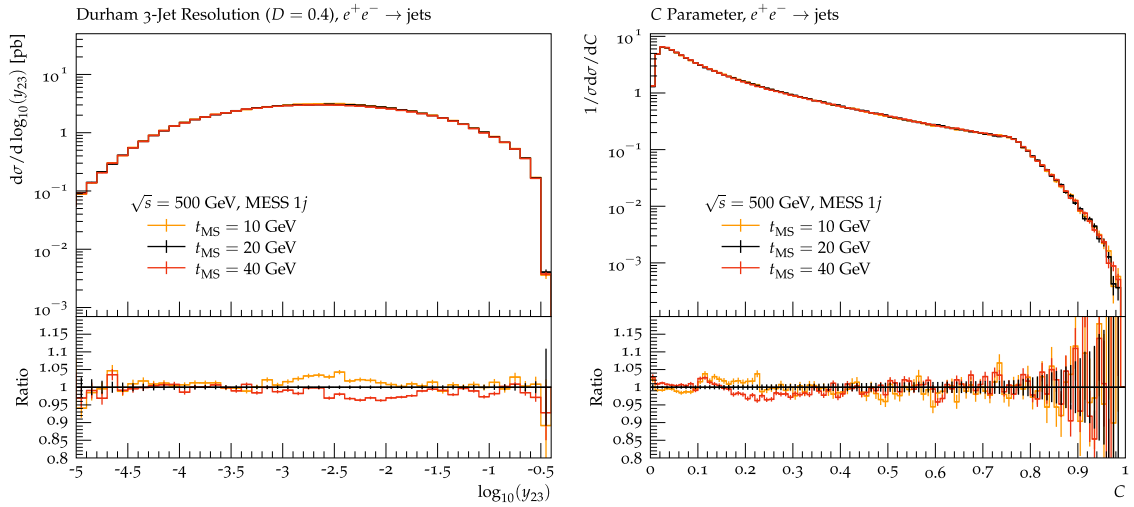


Fig. 4. Influence of the merging scale choice on sector shower plus one-jet merged predictions of the Durham 3-jet (*left*) and the C parameter (*right*) distribution in $e^+e^- \to$ jets at $\sqrt{s} = 500$ GeV.

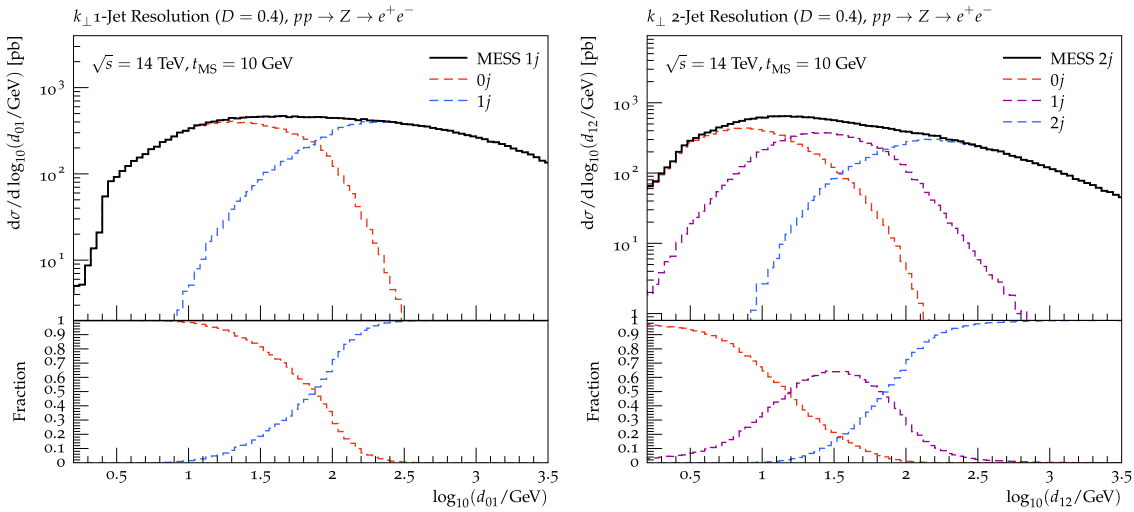


Fig. 5. Contributions of the individual hard-event samples in $k_\perp$ 1-jet (*left*) and 2-jet (*right*) resolution scales in the electron channel of Z production $pp \to Z +$ jets at $\sqrt{s} = 14$ TeV.

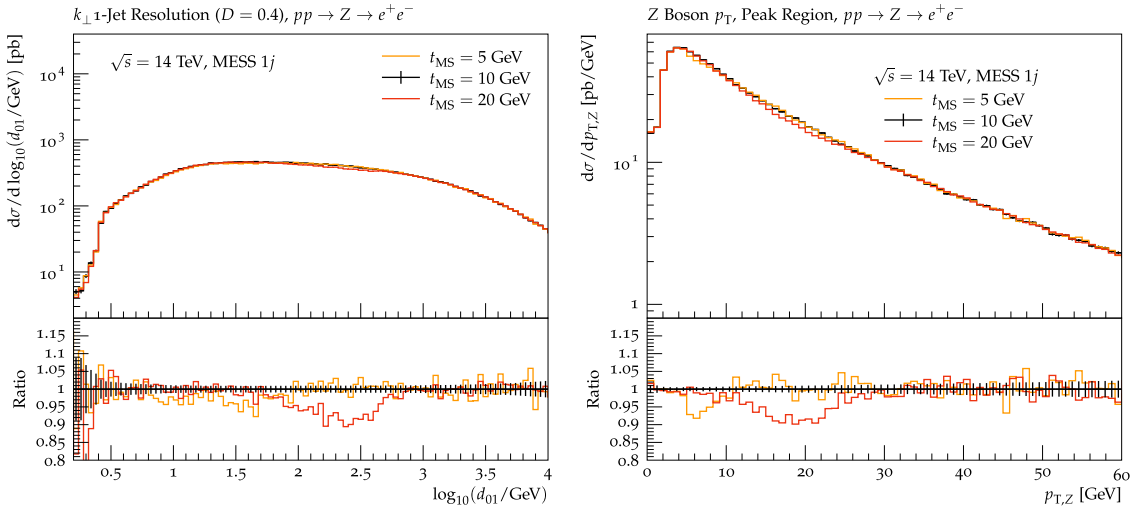


Fig. 6. Influence of the merging scale choice on sector shower plus one-jet merged predictions of the $k_\perp$ 1-jet (*left*) and the Z boson transverse momentum (*right*) distribution in $pp \to Z +$ jets at $\sqrt{s} = 14$ TeV.

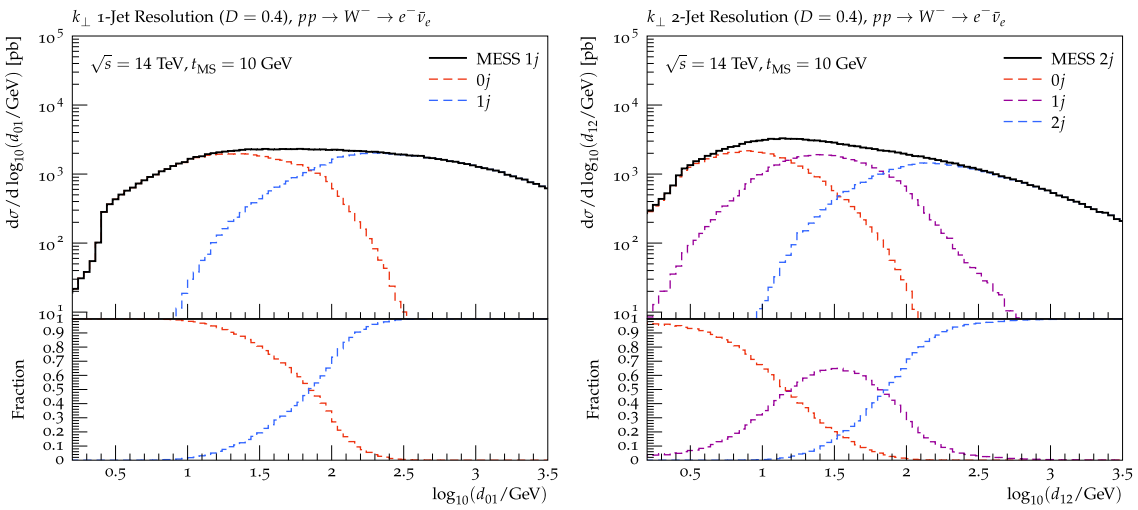


Fig. 7. Contributions of the individual hard-event samples in $k_\perp$ 1-jet (*left*) and 2-jet (*right*) resolution scales in the electron channel of W^- production $pp \to W^- +$ jets at $\sqrt{s} = 14$ TeV.

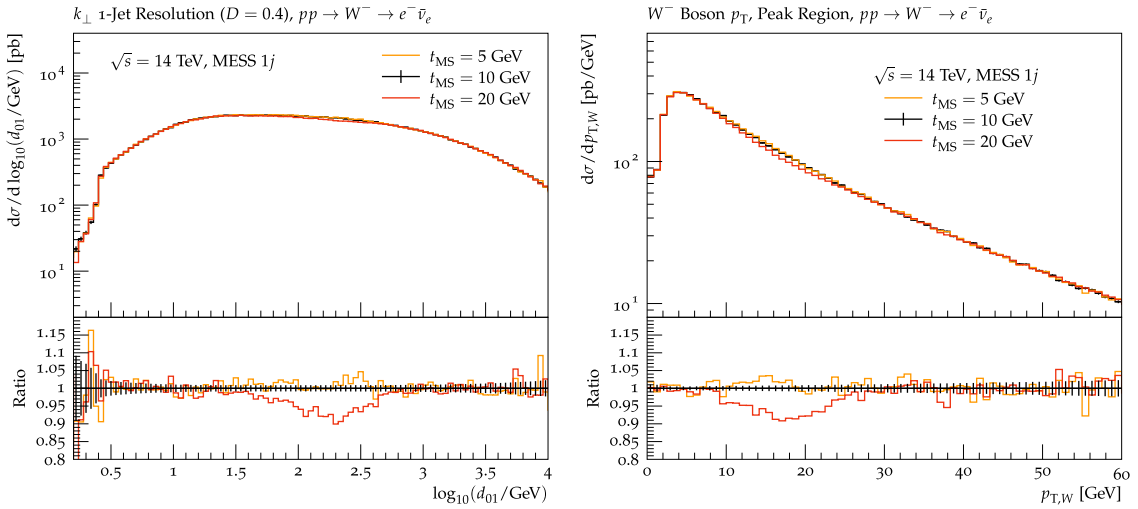


Fig. 8. Influence of the merging scale choice on sector shower plus one-jet merged predictions of the k_T 1-jet (*left*) and the W boson transverse momentum (*right*) distribution in $pp \to W^- +$ jets at $\sqrt{s} = 14$ TeV.

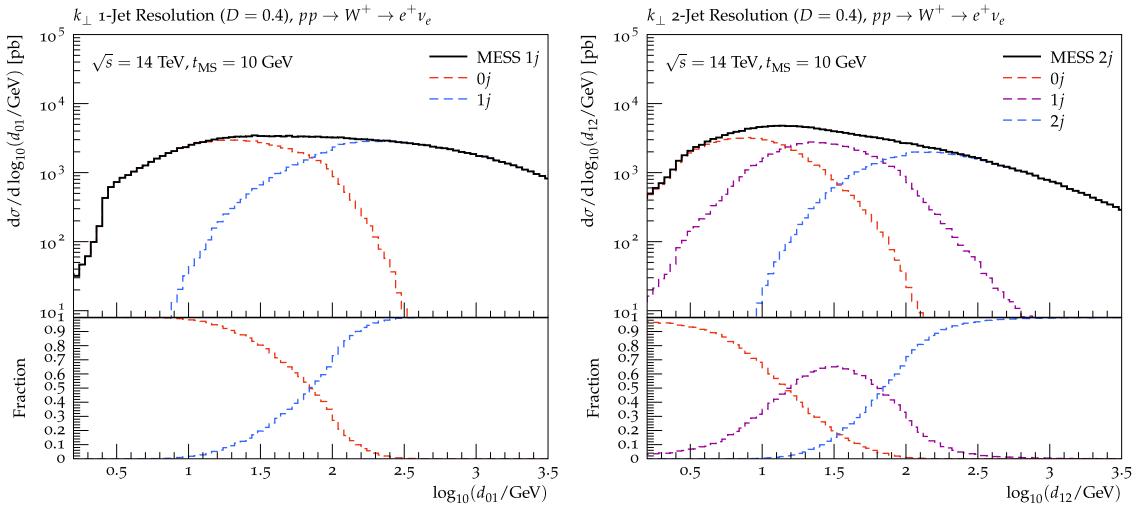


Fig. 9. Contributions of the individual hard-event samples in $k_\perp$ 1-jet (*left*) and 2-jet (*right*) resolution scales in the electron channel of W^+ production $pp \to W^+ +$ jets at $\sqrt{s} = 14$ TeV.

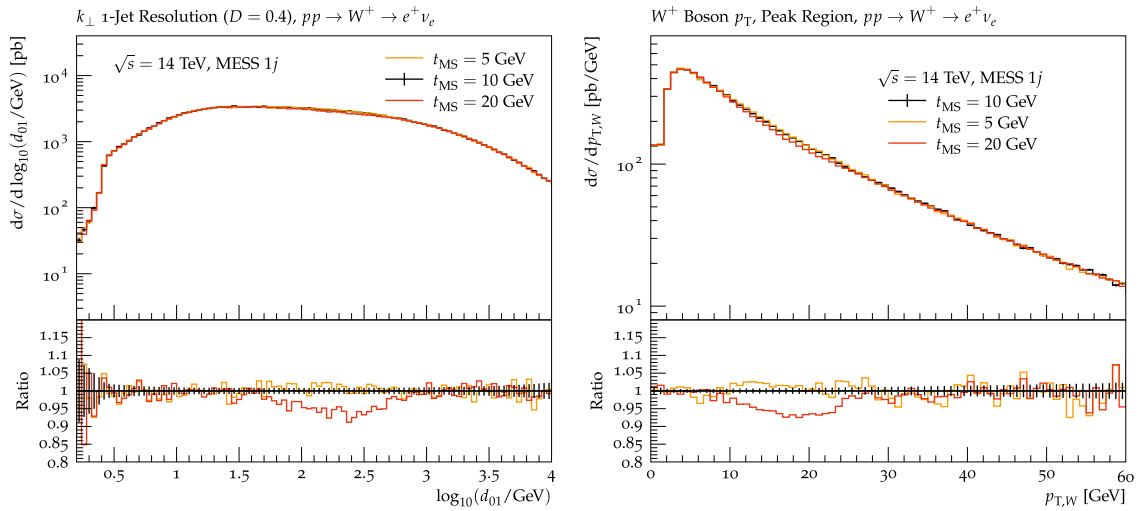


Fig. 10. Influence of the merging scale choice on sector shower plus one-jet merged predictions of the k_T 1-jet (*left*) and the W boson transverse momentum (*right*) distribution in $pp \to W^+ +$ jets at $\sqrt{s} = 14$ TeV.

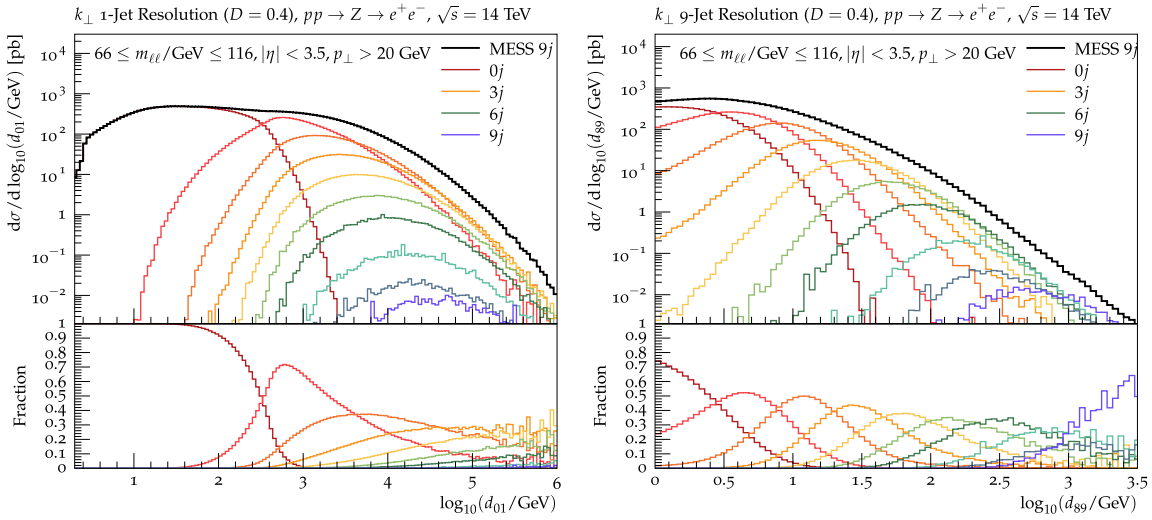


Fig. 11. VINCIA merged parton-level predictions for the $k_\perp$ 1-jet (*left*) and 9-jet (*right*) resolution scales in Drell-Yan processes in pp collisions at 14 TeV.

4. Results

To study the scaling behaviour of our implementation in the high-multiplicity regime, we use the parton-level event files [43–45] for vector boson production with up to 9 jets used in [42]. The merging scale is chosen to coincide with the $k_\perp$ cut of 20 GeV used in the event samples.

As a proof of concept, we show merged parton-level predictions for $k_\perp$ 1- and 9-jet resolution scales in $pp \to Z$ and $pp \to W^-$ with up to 9 hard jets in Figs. 11 and 12. These results are obtained with the preliminary default VINCIA tune with two-loop running α_s in the CMW scheme and $\alpha_s^{\overline{\rm MS}}(M_Z) = 0.118$ as summarised in Appendix A. Despite obtaining the accuracy of the additional tree-level matrix elements, the merged predictions retain the LL+LO precision of the shower, including unitarity violations due to the CKKW-L method. As events are weighted by α_s ratios, the non-unitarity of the method becomes manifest with increasing jet multiplicity, leading to larger scale dependencies at higher orders. To gain a first estimate of this effect, we vary VINCIA's renormalisation scale factors k_R used to evaluate the strong coupling (cf. Appendix A) by a factor of 2 with respect to the default values, cf. Fig. 13, where we also compare merged parton-level predictions for $k_\perp$ resolution scales in $pp \to W^-$ from VINCIA's and PYTHIA's CKKW-L implementation. For PYTHIA, we use the default tune of the strong coupling, i.e., a one-loop running coupling in the $\overline{\rm MS}$ scheme with a numerical value of $\alpha_s^{\overline{\rm MS}}(M_Z) = 0.1365$.

Except for the region near the hadronisation cutoff, the two implementations agree well for the 1-jet resolution scale, cf. the left-hand pane in Fig. 13, while there is a larger discrepancy in the distributions of the 9-jet resolution scale. As expected, the scale variations only have a small effect on the former, while for the 9-jet clustering scale, these are far more significant. There remains, however, a shape difference between VINCIA and PYTHIA for the 9-jet scale, which may at least partly be traced back to the rather high merging scale for the uncorrected sector shower, cf. Section 3. Given that neither of the showers are tuned for merging at these high multiplicities, the observed difference provides an interesting subject for further studies.

As the central objective of our improved merging scheme, we study the run time and memory usage of our implementation and compare it to the CKKW-L implementation in PYTHIA.

4.1. Run time

As a first gauge of the scaling behaviour of the default PYTHIA and our VINCIA sector shower CKKW-L implementations, we mea-

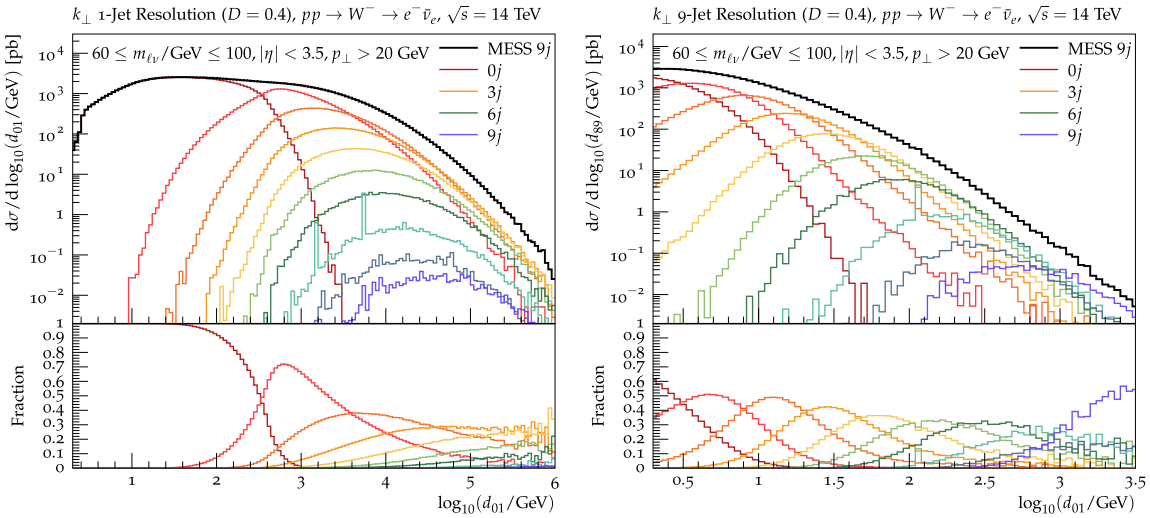


Fig. 12. VINCIA merged parton-level predictions for the $k_\perp$ 1-jet (*left*) and 9-jet (*right*) resolution scales in W^- production in pp collisions at 14 TeV.

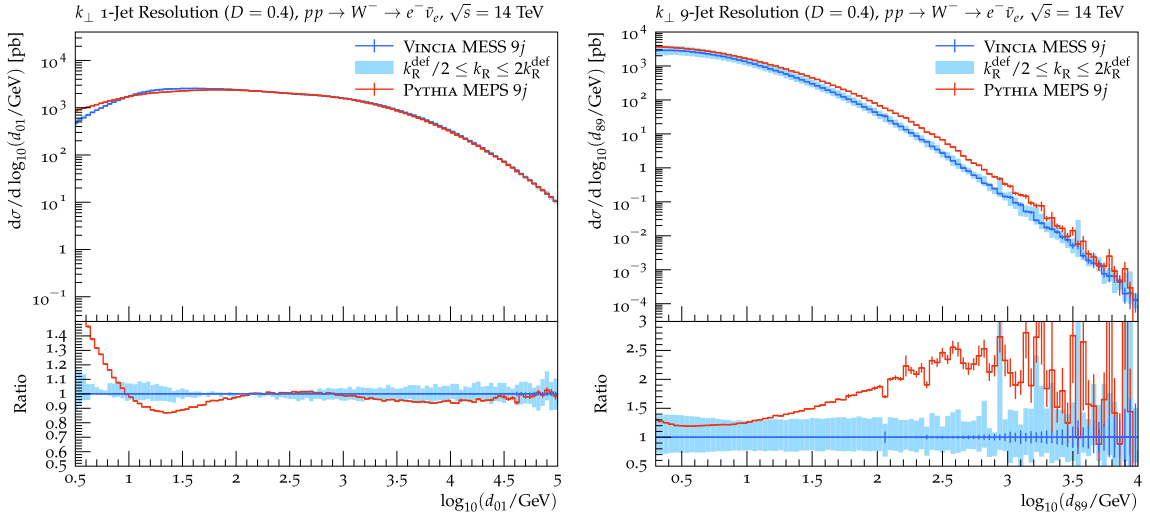


Fig. 13. Comparison of VINCIA and PYTHIA merged predictions for $k_\perp$ 1-jet (*left*) and 9-jet (*right*) resolution scales in $pp \to W^- +$ jets at $\sqrt{s} = 14$ TeV.

sure the CPU time to find the (most probable) shower history in both. We consider the processes $pp \to Z$ and $pp \to W^-$ with up to 9 additional hard jets and run each multiplicity individually on a single core of a 2.3 GHz Intel Core i5 processor and only count complete histories, i.e., ones for which at least one reconstructed shower sequence to the Born exists.

In the left-hand panes of Figs. 14 and 15, we show the scaling of the CPU time for shower history construction in $pp \to Z +$ jets and $pp \to W^- +$ jets, respectively. We find that the recursive strategy of the default PYTHIA history construction is faster for low multiplicities, but develops a steep exponential scaling for higher multiplicities. The iterated VINCIA sector shower history construction, on the other hand, scales linearly with the number of jets. Starting from the four-jet sample, it becomes notably faster than the PYTHIA history construction. At the extreme of 9 jets, PYTHIA spends about half a second per event to construct all shower histories, while VINCIA does not even need a millisecond per history.

As being of practical importance, we study the overall CPU time per generated event in the right-hand panes of Figs. 14 and 15. We consider the time for PYTHIA to generate a new parton-level event, either using the default merging and shower implementation or the VINCIA one. Again, the default implementation is notably faster for low multiplicities but develops a steep exponential scaling. By comparing with the time needed for history construction, it can be seen that, starting from the four-jet sample, the default PYTHIA implementation spends most of the event generation time on constructing the history of the input event. Because of the more complex shower algorithm, the picture is completely different for the VINCIA MESS implementation, where most of the time is spent on the showering of an accepted hard event. As the number of accepted events decreases with the multiplicity, and therefore less often a full sector shower has to be performed, the overall event generation time stays approximately constant when adding more jets, with a slight decrease towards high multiplicities. The increase for the 9-jet sample is explained by the inclusive treatment of the last node in CKKW-L merging, due to which a full sector shower is performed for more events. When adding further exclusive multiplicities beyond 8 jets, the total event generation time of the VINCIA MESS implementation will approach the pure history construction time.

It should be noted that the baseline sector shower used in this study does not utilise optimised sampling strategies to deal with competing sectors, cf. e.g. [68–70], which can improve the performance relative to the results shown here. Such optimisation studies are currently ongoing.

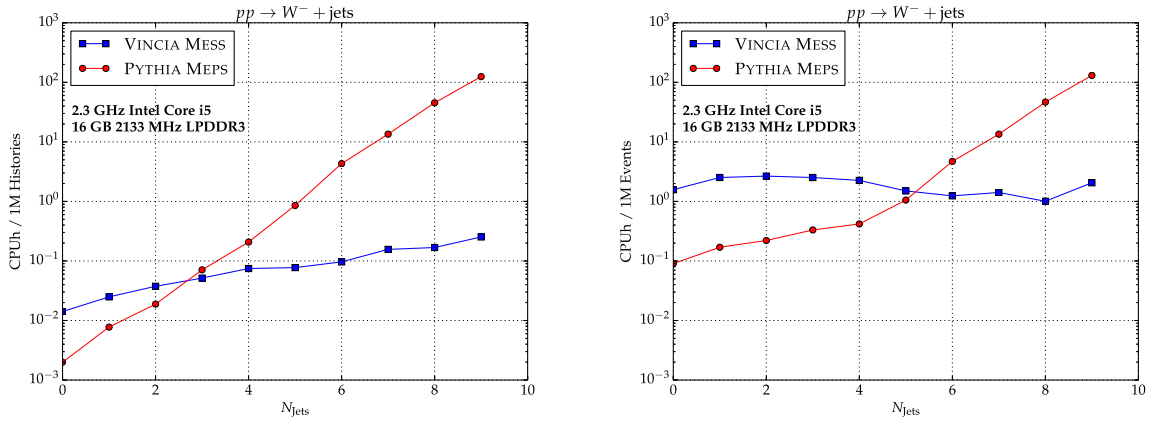


Fig. 14. PYTHIA and VINCIA CPU time scaling in history construction (*left*) and parton-level event generation (*right*) for $pp \to W^- + \text{jets}$ merging at $\sqrt{s} = 14$ TeV.

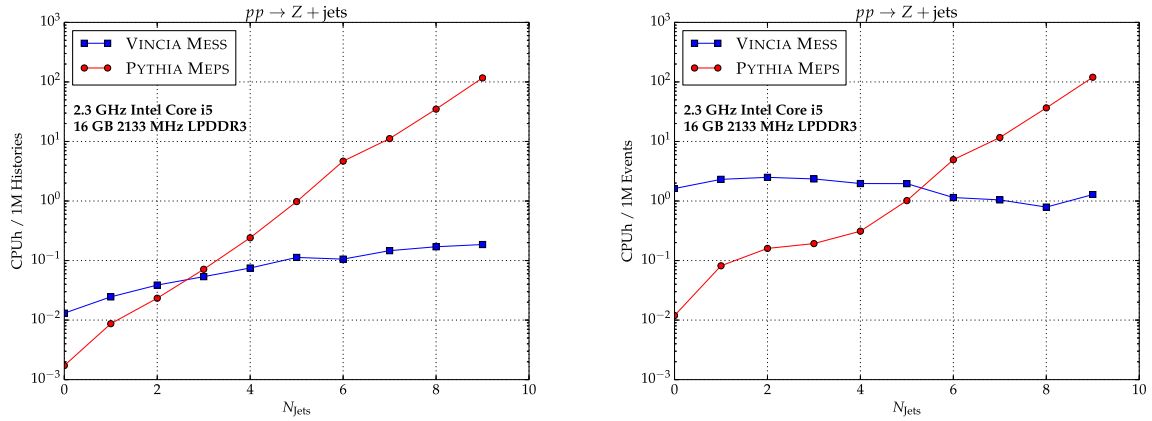


Fig. 15. PYTHIA and VINCIA CPU time scaling in history construction (*left*) and parton-level event generation (*right*) for $pp \to Z + \text{jets}$ merging at $\sqrt{s} = 14$ TeV.

4.2. Memory usage

As the even more prohibiting bottleneck of conventional CKKW-L merging schemes at high multiplicities, we study the memory usage. We use `Valgrind`'s `Massif` tool to monitor the heap usage of the default PYTHIA CKKW-L merging and our VINCIA sector shower merging implementations. In particular, this means that neither the stack nor the memory at the page level is recorded. For comparability and reproducibility, we use the `--time-unit=B` option in `Valgrind` to measure the runtime of the program in terms of the number of allocated and deallocated bytes. We use the same main program and event samples for both runs and consider a fictitious $Z + 10$ jet merging run, so that every event multiplicity, including the 9-jet sample, is processed as an intermediate node. We run each multiplicity independently with the maximal possible number of snapshots available, which may be at most (but is not necessarily identical to) 1000. To gain the most detailed possible picture of the memory allocations, we choose a relatively small number of 1000 events per run. For higher statistical significance, we perform up to ten independent runs for each multiplicity. On the technical level, the 7-, 8-, and 9-jet event samples in [45,44,43] are separated into multiple files, corresponding to irreducible groups of processes with similar diagrammatic structure, cf. [42]. For these multiplicities, at least one run is performed per group.

In Fig. 16, the individual heap profiles of all event samples from the $pp \to Z$ to the $pp \to Z + 9$ jets sample are shown. For samples with more than six jets, we only show a representative memory profile of one group. Additional profiles are collected in Appendix B. The peak on the left-hand side of the plots corresponds to the read-in of the HDF5 event sample, which (for high-multiplicity runs) is not recorded by `Valgrind` for the default CKKW-L implementation in PYTHIA anymore, in favour of higher allocation peaks later in the run. For merging with less than 4 additional jets, the heap profiles of PYTHIA and VINCIA are very similar: after the high peak when reading the event file, only a number of very small peaks are recorded. For these runs, the PYTHIA merging implementation has a shorter "runtime" in terms of total allocated/deallocated memory, which, however, continuously increases with the number of additional jets. That VINCIA allocates more memory than PYTHIA during these runs can be traced back to the differences in the shower implementations, which, as alluded to above, is more complicated for the sector shower. For every trial, a tentative post-branching state has to be constructed to evaluate the sector veto on. Although this does not amount to large peaks in memory allocation, it adds to the total allocated memory, i.e., the "run time". Beginning with the $Z + 5$ jets sample, sizeable effects become visible in the PYTHIA memory profiles. The peak heights continuously grow for PYTHIA and eventually outgrow the first file-reading peak. The VINCIA memory profile remains mostly constant and becomes negligible in comparison to PYTHIA's profile for the 8- and 9-jet samples. PYTHIA's history-construction technique directly translates to the memory profiles; after a history has been chosen, the memory allocation returns to the baseline value at which VINCIA remains throughout.

As a gauge of the scaling behaviour of the memory usage in both merging implementations, we plot the total allocated/deallocated memory per 1k events in Fig. 17. For each multiplicity, we average over statistically independent runs and from 7 jets on, we also average over the different groupings. While PYTHIA shows a

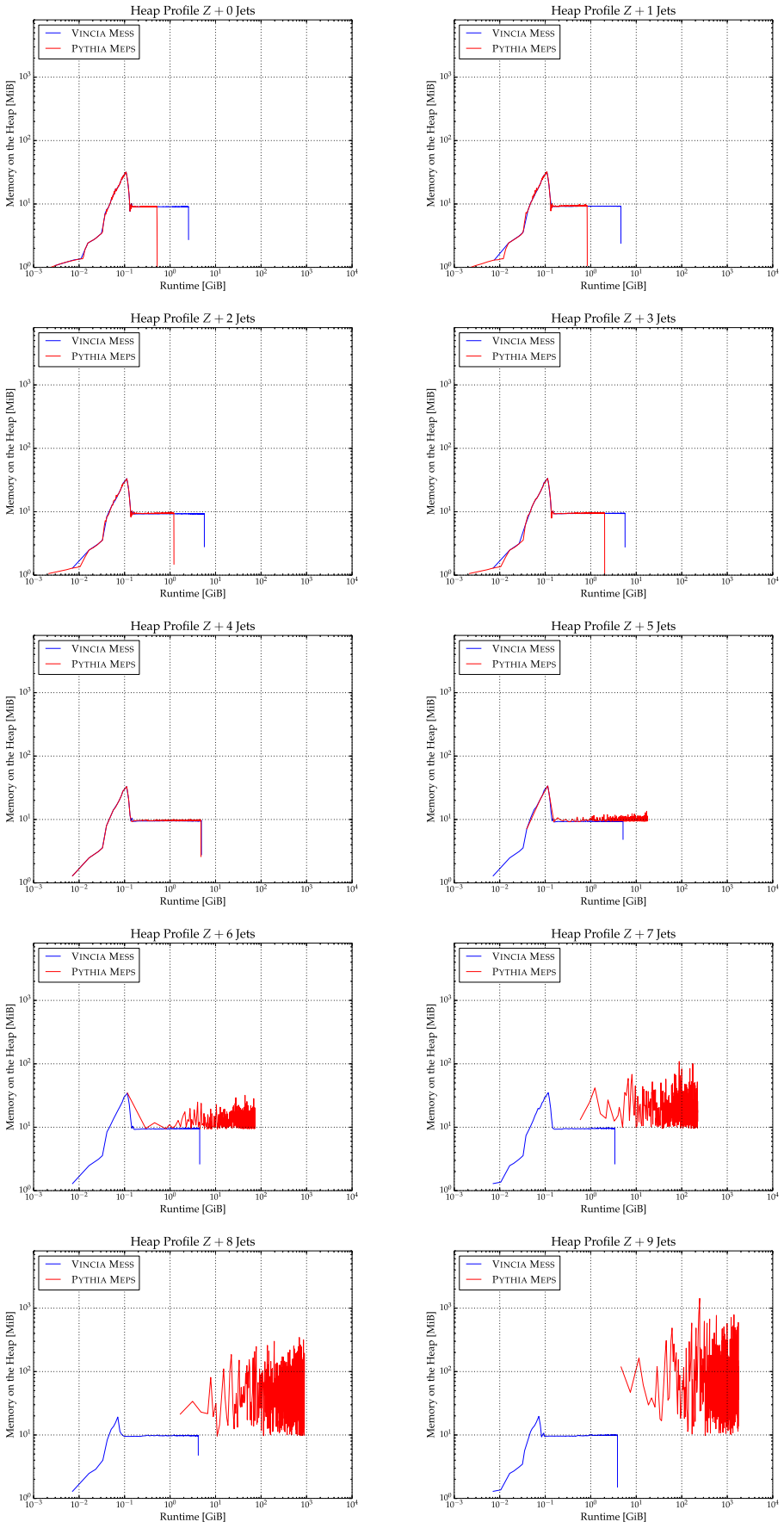


Fig. 16. PYTHIA and VINCIA memory usage profiles in $pp \to Z +$ jets merging at $\sqrt{s} = 14$ TeV.

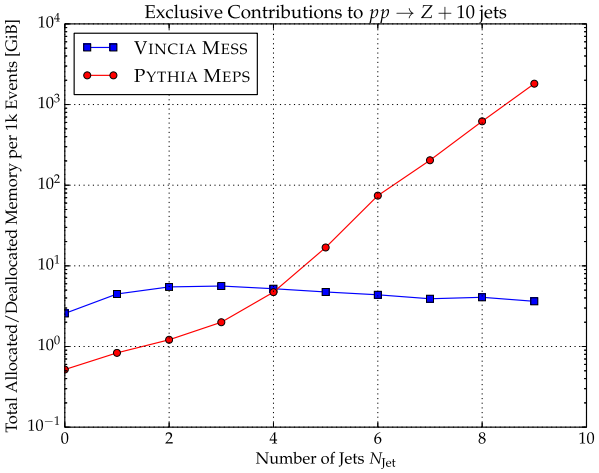


Fig. 17. PYTHIA and VINCIA memory usage scaling in $pp \to Z +$ jets merging at $\sqrt{s} =$ 14 TeV.

rather dramatic scaling, with allocating and deallocating a total of 1 TiB of data for $Z + 9$ jets, the VINCIA curve remains almost flat, with only a small peak around 3 additional jets. The latter can be understood by considering that the sector shower has a comparable memory footprint as the merging and that in the latter maximally two histories are stored concurrently, cf. Section 2.3. At high multiplicities, most of the events get vetoed during the trial showers and the sector shower is never started off these events. For samples with 1 – 3 additional jets, on the other hand, a fair number of events are accepted and further processed by the sector shower, explaining the small increase in memory usage there.

5. Conclusions

We here presented the first-ever implementation of the CKKW-L merging approach with sector showers, which alleviates the bottlenecks of conventional implementations while accurately calculating the Sudakov factors as generated by the shower. The merging scheme was implemented for the VINCIA antenna shower in the PYTHIA 8.3 event generator; this implementation is mostly independent from the default CKKW-L one, and has been made public in the PYTHIA 8.304 release.

We have validated the implementation for processes of immediate phenomenological interest and studied the scaling behaviour of the method in multi-jet merging in vector boson production at high multiplicities. While the time to construct sector shower histories scales approximately linearly with the number of hard jets, the overall event generation time as well as the memory usage stays approximately constant. Both provides a significant improvement over the exponential scaling of the default merging implementation in PYTHIA. As a consequence, including merging hard jets with the sector shower in fact becomes easier with increasing multiplicity. We gained a first estimate of renormalisation scale uncertainties arising at high merged multiplicities and compared preliminary results to PYTHIA's CKKW-L implementation.

While we have here focused on the computational improvements, a dedicated physics study with the MESS framework is yet to follow. In such a study, the default VINCIA tune should be reviewed to achieve an accurate description on the hadron level when including higher-order matrix elements. Moreover, immediate future work can be done on improving sampling methods for sector antennae in their respective sectors, which is currently only inefficiently achieved by means of multiple global antennae over their full phase spaces and rejecting branchings outside of appropriate sectors. This results in the overall slower shower algorithm, as seen above.

As the sector merging approach developed here is an adaption of the CKKW-L technique, existing refinements and extensions of it can readily be applied. As such, it can be extended to retain unitarity by the methods presented in [16,26] or to include NLO corrections by the methods of [19,17,18,22,23] in the future. Even extensions towards NNLO are feasible along the lines of the method presented in [29–31]. For these, only the construction of the shower histories needs to be adapted to the sector case. To this end, the NL3 [19] and UNLOPS [17] schemes are particularly well suited for the generalisation to NLO merging, as their adaption to sector showers would follow the exact same steps as in the CKKW-L case outlined in Section 2. This means that only the probabilistic construction of shower histories would be replaced by the sectorised history construction, while all other steps in NL3 and UNLOPS stay the same.

In a less straightforward way, it might also be possible to extend the LO merging technique presented here to higher orders via other schemes, such as the MINLO [24] and MINNLO$_{\text{PS}}$ [33] or FxFx [25] techniques. These extensions do, however, differ in the way histories are constructed and scales are associated to intermediate states. In addition to an adaptation of the history construction to sector showers, an implementation of these schemes would, for instance, also require the implementation of analytical Sudakov factors.

In the light of the observed scale uncertainties at high multiplicities, both unitarity-improved schemes and extensions to the NLO (or even higher orders) provide interesting and sensible avenues for future work. It is worth pointing out that in both cases, event files are usually processed twice to generate counter terms. As the algorithm presented here will always yield a unique history for an input configuration, it bears the potential to make both unitary and NLO merging even more efficient, as event files might only have to be processed once. On the one hand, this reduces the overall run time and on the other hand, this might alleviate problems regarding negative-weight events. As the implementation in the VINCIA shower furthermore implies a dedicated interleaved resonance shower framework [71,70,72], it may be worthwhile to explore merging in coloured as well as electroweak resonance systems in the future.

Although the discussion here was tailored to the implementation in VINCIA, it demonstrates the efficiency gains that can be obtained through a relatively straightforward adaption of the shower model. We have sketched how our method could be adopted in other shower models and exemplified it for the case of Catani-Seymour dipole showers in Section 2.2.

We close by noting that with the merging scheme presented here, shower-plus-matrix-element calculations with more than 9 hard jets are readily possible on the shower side. The main bottlenecks of merged calculations remain entirely on the fixed-order side and generating large numbers of high-multiplicity configurations still remains a time- and resource-consuming endeavour.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Peter Skands for many helpful and fruitful discussions and comments on the manuscript. CTP thanks Stefan Höche for help with the HDF5 event files and Stefan Prestel for help with the HDF5 data format and clarifications on the CKKW-L implementation in PYTHIA. CTP also thanks Steffen Schumann and Enrico Bothmann for helpful comments on commen-

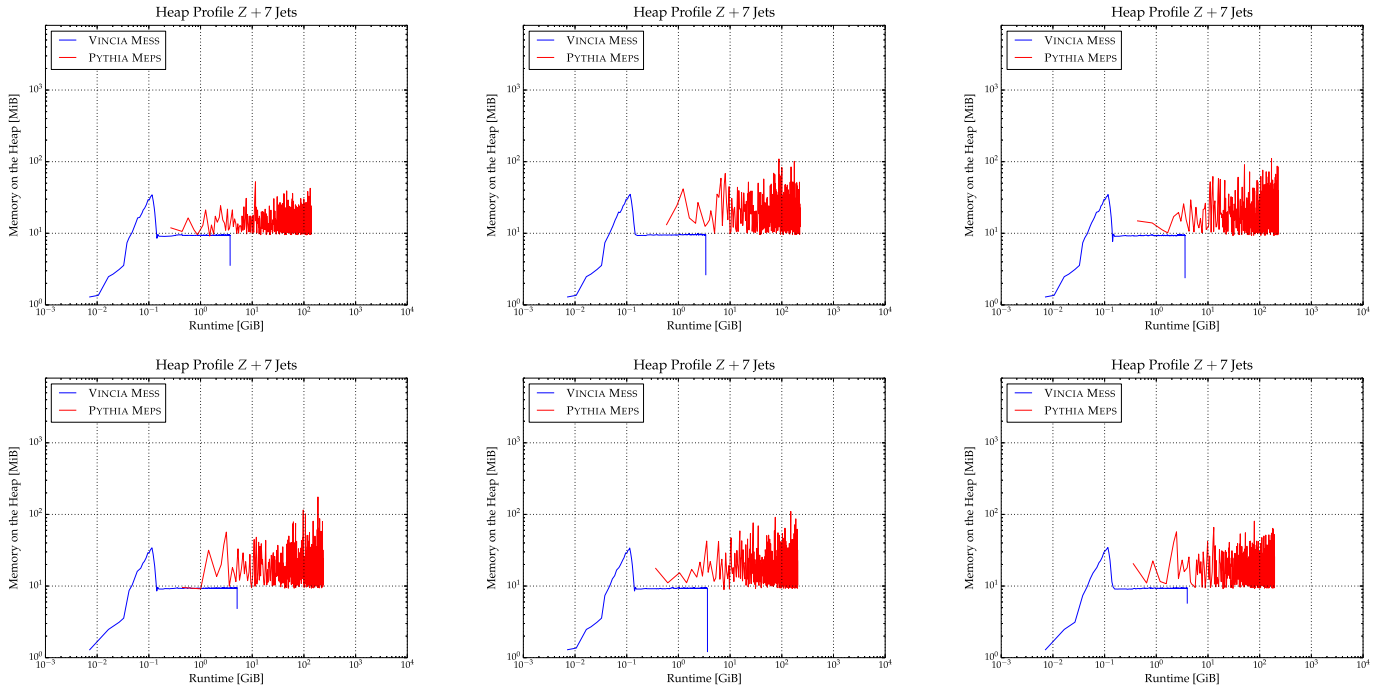


Fig. B.18. PYTHIA and VINCIA memory usage profiles for different process groupings in $pp \to Z + 7$ jets samples at $\sqrt{s} = 14$ TeV.

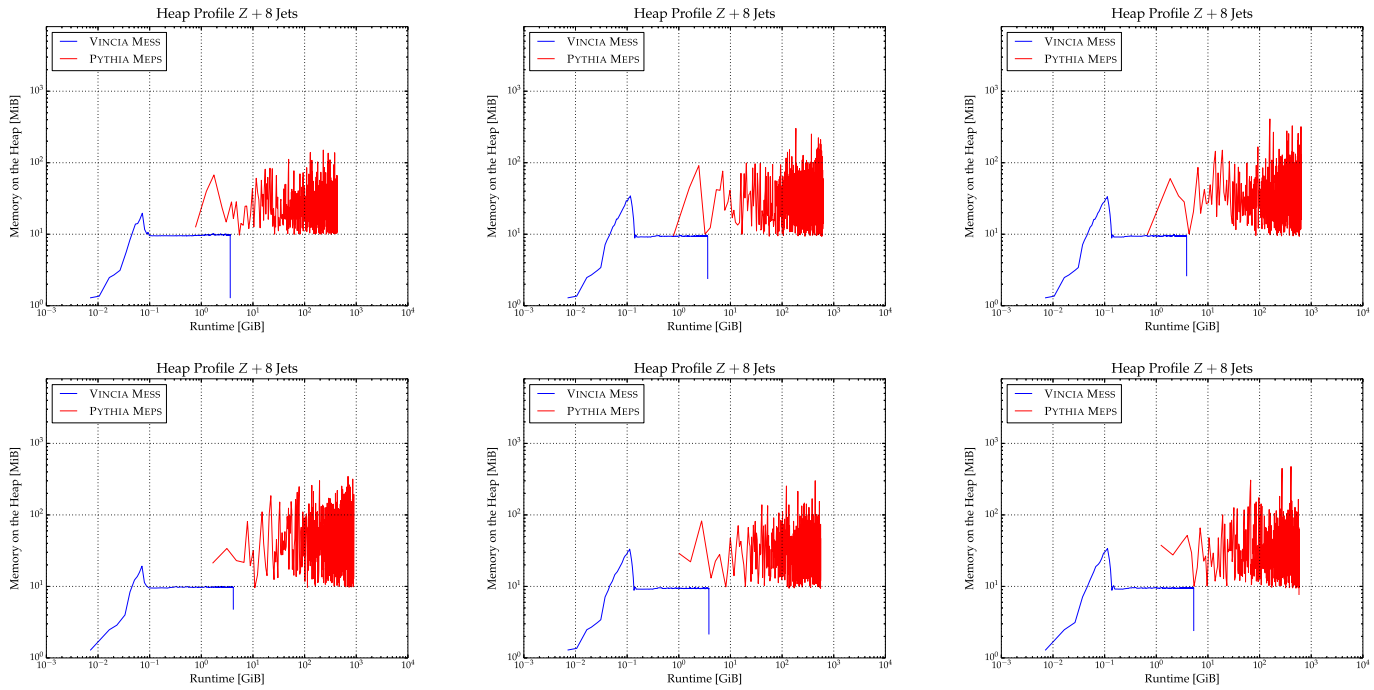


Fig. B.19. PYTHIA and VINCIA memory usage profiles for different process groupings in $pp \to Z + 8$ jets samples at $\sqrt{s} = 14$ TeV.

surate sector scales. We thank Cody B Duncan for proofreading the manuscript. We acknowledge support from the Monash eResearch Centre and eSolutions-Research Support Services through the MonARCH HPC Cluster. CTP is supported by the Monash Graduate Scholarship, the Monash International Postgraduate Research Scholarship, and the J. L. William Scholarship. HB received funding from the Australian Research Council via Discovery Project DP170100708 – "Emergent Phenomena in Quantum Chromodynamics". This work was also supported in part by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 722104 – MCnetITN3.

Appendix A. Perturbative tune parameters

We use the preliminary default VINCIA tune of the perturbative parameters, cf. [47]. We include two-loop running-coupling effects with an effective value of α_s chosen according to the CMW scheme [73],

$$\alpha_s^{\mathrm{CMW}} = \alpha_s^{\overline{\mathrm{MS}}}\left(1 + \frac{\alpha_s^{\overline{\mathrm{MS}}}}{2\pi}\left[C_A\left(\frac{67}{18} - \frac{\pi^2}{6}\right) - \frac{5n_f}{9}\right]\right),$$

$$\alpha_s^{\overline{\mathrm{MS}}}(M_Z) = 0.118\,, \tag{A.1}$$

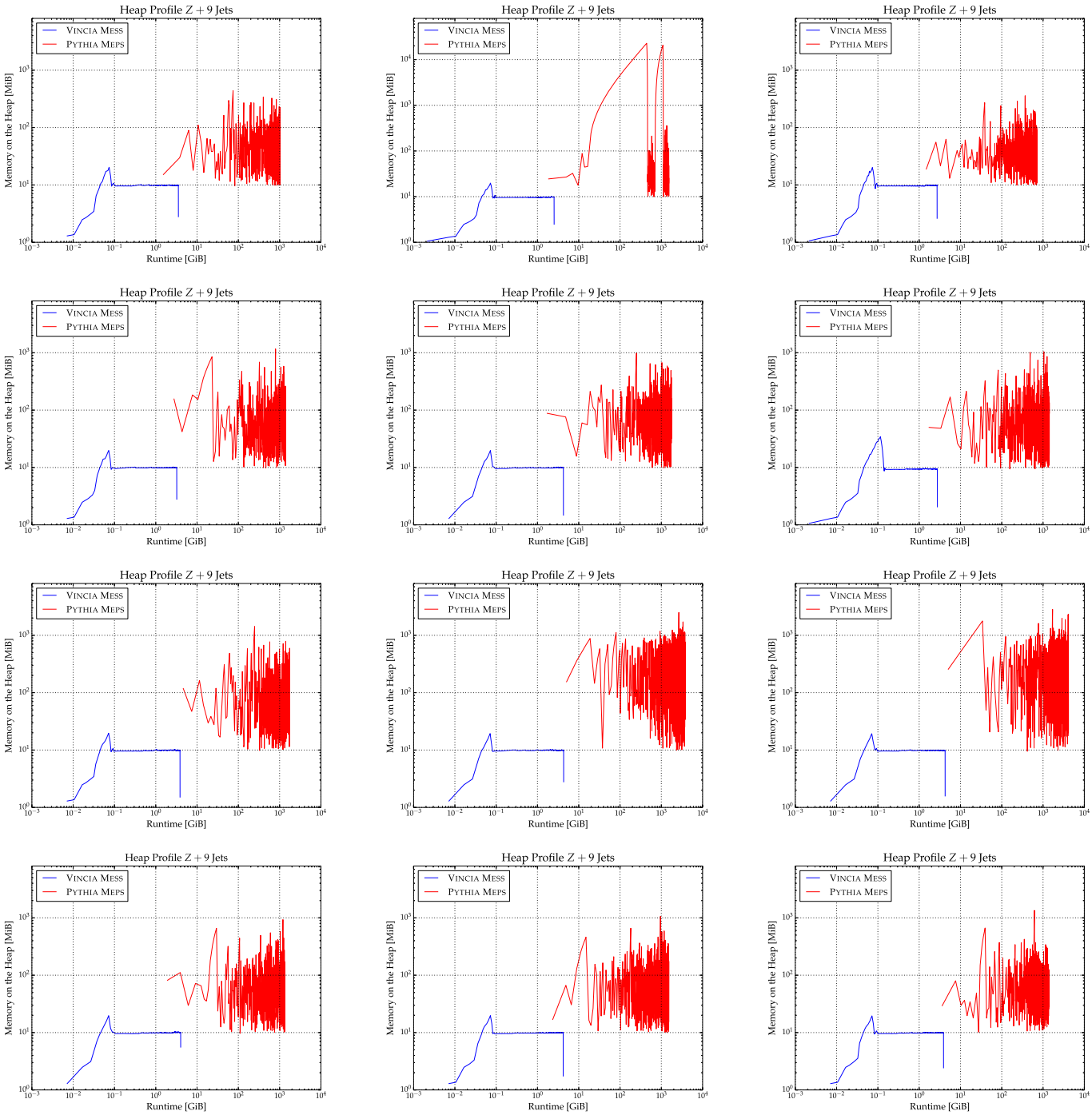


Fig. B.20. Pythia and Vincia memory usage profiles for different process groupings in $pp \to Z + 9$ jets samples at $\sqrt{s} = 14$ TeV.

supplemented by renormalisation-scale prefactors k_R, which modify the evolution scale, $p_\perp$, in the argument of the running coupling,

$$\alpha_s^{\text{Vincia}}(p_\perp^2) = \alpha_s^{\text{CMW}}(k_R\, p_\perp^2)\,. \tag{A.2}$$

The default values for these additional scale prefactors are chosen based on preliminary studies of LEP event shapes and Drell-Yan $p_\perp$-spectra as

$$k^{\text{F}}_{\text{R,Emit}} = 0.66\,, \quad k^{\text{F}}_{\text{R,Split}} = 0.8\,, \tag{A.3}$$

$$k^{\text{I}}_{\text{R,Emit}} = 0.66\,, \quad k^{\text{I}}_{\text{R,Split}} = 0.5\,, \quad k^{\text{I}}_{\text{R,Conv}} = 0.5\,. \tag{A.4}$$

Appendix B. Memory profiles

We here collect additional memory profiles for event samples with more than 6 jets, for which the event files are grouped according to similar process structures, cf. Figs. B.18, B.19, and B.20.

References

[1] G. Aad, et al., Eur. Phys. J. C 75 (2) (2015) 82, https://doi.org/10.1140/epjc/s10052-015-3262-7, arXiv:1409.8639.
[2] M. Aaboud, et al., Eur. Phys. J. C 77 (6) (2017) 361, https://doi.org/10.1140/epjc/s10052-017-4900-z, arXiv:1702.05725.
[3] M. Bengtsson, T. Sjöstrand, Phys. Lett. B 185 (1987) 435, https://doi.org/10.1016/0370-2693(87)91031-8.

[4] W.T. Giele, D.A. Kosower, P.Z. Skands, Phys. Rev. D 78 (2008) 014026, https://doi.org/10.1103/PhysRevD.78.014026, arXiv:0707.3652.
[5] W.T. Giele, D.A. Kosower, P.Z. Skands, Phys. Rev. D 84 (2011) 054003, https://doi.org/10.1103/PhysRevD.84.054003, arXiv:1102.2126.
[6] N. Fischer, S. Prestel, Eur. Phys. J. C 77 (9) (2017) 601, https://doi.org/10.1140/epjc/s10052-017-5160-7, arXiv:1706.06218.
[7] S. Frixione, B.R. Webber, J. High Energy Phys. 06 (2002) 029, https://doi.org/10.1088/1126-6708/2002/06/029, arXiv:hep-ph/0204244.
[8] P. Nason, J. High Energy Phys. 11 (2004) 040, https://doi.org/10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
[9] S. Frixione, P. Nason, C. Oleari, J. High Energy Phys. 11 (2007) 070, https://doi.org/10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
[10] M.L. Mangano, M. Moretti, R. Pittau, Nucl. Phys. B 632 (2002) 343–362, https://doi.org/10.1016/S0550-3213(02)00249-3, arXiv:hep-ph/0108069.
[11] M.L. Mangano, M. Moretti, F. Piccinini, M. Treccani, J. High Energy Phys. 01 (2007) 013, https://doi.org/10.1088/1126-6708/2007/01/013, arXiv:hep-ph/0611129.
[12] S. Catani, F. Krauss, R. Kuhn, B.R. Webber, J. High Energy Phys. 11 (2001) 063, https://doi.org/10.1088/1126-6708/2001/11/063, arXiv:hep-ph/0109231.
[13] L. Lönnblad, J. High Energy Phys. 05 (2002) 046, https://doi.org/10.1088/1126-6708/2002/05/046, arXiv:hep-ph/0112284.
[14] S. Hoeche, F. Krauss, S. Schumann, F. Siegert, J. High Energy Phys. 05 (2009) 053, https://doi.org/10.1088/1126-6708/2009/05/053, arXiv:0903.1219.
[15] K. Hamilton, P. Richardson, J. Tully, J. High Energy Phys. 11 (2009) 038, https://doi.org/10.1088/1126-6708/2009/11/038, arXiv:0905.3072.
[16] L. Lönnblad, S. Prestel, J. High Energy Phys. 03 (2012) 019, https://doi.org/10.1007/JHEP03(2012)019, arXiv:1109.4829.
[17] L. Lönnblad, S. Prestel, J. High Energy Phys. 02 (2013) 094, https://doi.org/10.1007/JHEP02(2013)094, arXiv:1211.4827.
[18] S. Plätzer, J. High Energy Phys. 08 (2013) 114, https://doi.org/10.1007/JHEP08(2013)114, arXiv:1211.5467.
[19] N. Lavesson, L. Lönnblad, J. High Energy Phys. 12 (2008) 070, https://doi.org/10.1088/1126-6708/2008/12/070, arXiv:0811.2912.
[20] S. Höche, F. Krauss, M. Schönherr, F. Siegert, J. High Energy Phys. 08 (2011) 123, https://doi.org/10.1007/JHEP08(2011)123, arXiv:1009.1127.
[21] L. Lönnblad, S. Prestel, J. High Energy Phys. 03 (2013) 166, https://doi.org/10.1007/JHEP03(2013)166, arXiv:1211.7278.
[22] T. Gehrmann, S. Höche, F. Krauss, M. Schönherr, F. Siegert, J. High Energy Phys. 01 (2013) 144, https://doi.org/10.1007/JHEP01(2013)144, arXiv:1207.5031.
[23] S. Hoeche, F. Krauss, M. Schönherr, F. Siegert, J. High Energy Phys. 04 (2013) 027, https://doi.org/10.1007/JHEP04(2013)027, arXiv:1207.5030.
[24] K. Hamilton, P. Nason, G. Zanderighi, J. High Energy Phys. 10 (2012) 155, https://doi.org/10.1007/JHEP10(2012)155, arXiv:1206.3572.
[25] R. Frederix, S. Frixione, J. High Energy Phys. 12 (2012) 061, https://doi.org/10.1007/JHEP12(2012)061, arXiv:1209.6215.
[26] J. Bellm, S. Gieseke, S. Plätzer, Eur. Phys. J. C 78 (3) (2018) 244, https://doi.org/10.1140/epjc/s10052-018-5723-2, arXiv:1705.06700.
[27] K. Hamilton, P. Nason, C. Oleari, G. Zanderighi, J. High Energy Phys. 05 (2013) 082, https://doi.org/10.1007/JHEP05(2013)082, arXiv:1212.4504.
[28] S. Alioli, C.W. Bauer, C. Berggren, F.J. Tackmann, J.R. Walsh, Phys. Rev. D 92 (9) (2015) 094020, https://doi.org/10.1103/PhysRevD.92.094020, arXiv:1508.01475.
[29] S. Höche, Y. Li, S. Prestel, Phys. Rev. D 91 (7) (2015) 074015, https://doi.org/10.1103/PhysRevD.91.074015, arXiv:1405.3607.
[30] S. Höche, Y. Li, S. Prestel, Phys. Rev. D 90 (5) (2014) 054011, https://doi.org/10.1103/PhysRevD.90.054011, arXiv:1407.3773.
[31] S. Höche, S. Kuttimalai, Y. Li, Phys. Rev. D 98 (11) (2018) 114013, https://doi.org/10.1103/PhysRevD.98.114013, arXiv:1809.04192.
[32] S. Alioli, A. Broggio, S. Kallweit, M.A. Lim, L. Rottoli, Phys. Rev. D 100 (9) (2019) 096016, https://doi.org/10.1103/PhysRevD.100.096016, arXiv:1909.02026.
[33] P.F. Monni, P. Nason, E. Re, M. Wiesemann, G. Zanderighi, J. High Energy Phys. 05 (2020) 143, https://doi.org/10.1007/JHEP05(2020)143, arXiv:1908.06987.
[34] C.G. Papadopoulos, Comput. Phys. Commun. 137 (2001) 247–254, https://doi.org/10.1016/S0010-4655(01)00163-1, arXiv:hep-ph/0007335.
[35] A. Kanaki, C.G. Papadopoulos, Comput. Phys. Commun. 132 (2000) 306–315, https://doi.org/10.1016/S0010-4655(00)00151-X, arXiv:hep-ph/0002082.
[36] F. Krauss, R. Kuhn, G. Soff, J. High Energy Phys. 02 (2002) 044, https://doi.org/10.1088/1126-6708/2002/02/044, arXiv:hep-ph/0109036.
[37] M. Moretti, T. Ohl, J. Reuter, in: T. Behnke, S. Bertolucci, R.D. Heuer, D. Miller, F. Richard, R. Settles, V. Telnov, P. Zerwas (Eds.), Proceedings, Physics and Experimentation at a Linear Electron-Positron Collider,2nd ECFA/DESY Study, 1998–2001, vols. 1–3, Lund, Sweden, June 28–30, 1998–2001, 2001, pp. 1981–2009, arXiv:hep-ph/0102195.
[38] M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A.D. Polosa, J. High Energy Phys. 07 (2003) 001, https://doi.org/10.1088/1126-6708/2003/07/001, arXiv:hep-ph/0206293.
[39] W. Kilian, T. Ohl, J. Reuter, Eur. Phys. J. C 71 (2011) 1742, https://doi.org/10.1140/epjc/s10052-011-1742-y, arXiv:0708.4233.
[40] T. Gleisberg, S. Hoeche, J. High Energy Phys. 12 (2008) 039, https://doi.org/10.1088/1126-6708/2008/12/039, arXiv:0808.3674.
[41] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.S. Shao, T. Stelzer, P. Torrielli, M. Zaro, J. High Energy Phys. 07 (2014) 079, https://doi.org/10.1007/JHEP07(2014)079, arXiv:1405.0301.
[42] S. Höche, S. Prestel, H. Schulz, Phys. Rev. D 100 (1) (2019) 014024, https://doi.org/10.1103/PhysRevD.100.014024, arXiv:1905.05120.
[43] H. Schulz, S. Höche, S. Prestel, Z + up to 9 jets parton level events at 14 TeV in HDF5, https://doi.org/10.5281/zenodo.2678039, May 2019.
[44] H. Schulz, S. Höche, S. Prestel, Wminus + up to 9 jets parton level events at 14 TeV in HDF5, https://doi.org/10.5281/zenodo.2678091, May 2019.
[45] H. Schulz, S. Höche, S. Prestel, Wplus + up to 9 jets parton level events at 14 TeV in HDF5, https://doi.org/10.5281/zenodo.2678055, May 2019.
[46] J.J. Lopez-Villarejo, P.Z. Skands, J. High Energy Phys. 11 (2011) 150, https://doi.org/10.1007/JHEP11(2011)150, arXiv:1109.3608.
[47] H. Brooks, C.T. Preuss, P. Skands, J. High Energy Phys. 07 (2020) 32, https://doi.org/10.1007/JHEP07(2020)032, arXiv:2003.00702.
[48] F. Krauss, J. High Energy Phys. 08 (2002) 015, https://doi.org/10.1088/1126-6708/2002/08/015, arXiv:hep-ph/0205283.
[49] N. Lavesson, L. Lönnblad, J. High Energy Phys. 04 (2008) 085, https://doi.org/10.1088/1126-6708/2008/04/085, arXiv:0712.2966.
[50] D.A. Kosower, Phys. Rev. D 57 (1998) 5410–5416, https://doi.org/10.1103/PhysRevD.57.5410, arXiv:hep-ph/9710213.
[51] D.A. Kosower, Phys. Rev. D 71 (2005) 045016, https://doi.org/10.1103/PhysRevD.71.045016, arXiv:hep-ph/0311272.
[52] Y.L. Dokshitzer, G. Marchesini, J. High Energy Phys. 03 (2009) 117, https://doi.org/10.1088/1126-6708/2009/03/117, arXiv:0809.1749.
[53] Z. Nagy, D.E. Soper, J. High Energy Phys. 05 (2009) 088, https://doi.org/10.1088/1126-6708/2009/05/088, arXiv:0901.3587.
[54] P.Z. Skands, S. Weinzierl, Phys. Rev. D 79 (2009) 074021, https://doi.org/10.1103/PhysRevD.79.074021, arXiv:0903.2150.
[55] G. Bewick, S. Ferrario Ravasio, P. Richardson, M.H. Seymour, J. High Energy Phys. 04 (2020) 019, https://doi.org/10.1007/JHEP04(2020)019, arXiv:1904.11866.
[56] M. Dasgupta, F.A. Dreyer, K. Hamilton, P.F. Monni, G.P. Salam, G. Soyez, Phys. Rev. Lett. 125 (5) (2020) 052002, https://doi.org/10.1103/PhysRevLett.125.052002, arXiv:2002.11114.
[57] C. Friberg, G. Gustafson, J. Hakkinen, Nucl. Phys. B 490 (1997) 289–305, https://doi.org/10.1016/S0550-3213(97)00064-3, arXiv:hep-ph/9604347.
[58] G. Gustafson, Nucl. Phys. B 392 (1993) 251–280, https://doi.org/10.1016/0550-3213(93)90203-2.
[59] M. Dasgupta, F.A. Dreyer, K. Hamilton, P.F. Monni, G.P. Salam, J. High Energy Phys. 09 (2018) 033, https://doi.org/10.1007/JHEP09(2018)033, arXiv:1805.09327.
[60] K. Hamilton, R. Medves, G.P. Salam, L. Scyboz, G. Soyez, Colour and logarithmic accuracy in final-state parton showers (11 2020), arXiv:2011.10054.
[61] J. Holguin, J.R. Forshaw, S. Plätzer, Improvements on dipole shower colour (11 2020), arXiv:2011.15087.
[62] J.R. Forshaw, J. Holguin, S. Plätzer, J. High Energy Phys. 09 (2020) 014, https://doi.org/10.1007/JHEP09(2020)014, arXiv:2003.06400.
[63] S. Catani, M. Seymour, Nucl. Phys. B 485 (1997) 291–419, https://doi.org/10.1016/S0550-3213(96)00589-5, arXiv:hep-ph/9605323, Erratum: Nucl. Phys. B 510 (1998) 503–504.
[64] S. Catani, S. Dittmaier, M.H. Seymour, Z. Trocsanyi, Nucl. Phys. B 627 (2002) 189–265, https://doi.org/10.1016/S0550-3213(02)00098-6, arXiv:hep-ph/0201036.
[65] A. Buckley, J. Butterworth, L. Lönnblad, D. Grellscheid, H. Hoeth, J. Monk, H. Schulz, F. Siegert, Comput. Phys. Commun. 184 (2013) 2803–2819, https://doi.org/10.1016/j.cpc.2013.05.021, arXiv:1003.0694.
[66] C. Bierlich, et al., SciPost Phys. 8 (2020) 026, https://doi.org/10.21468/SciPostPhys.8.2.026, arXiv:1912.05451.
[67] G. Parisi, Phys. Lett. B 74 (1978) 65–67, https://doi.org/10.1016/0370-2693(78)90061-8.
[68] R. Kleiss, R. Verheyen, Eur. Phys. J. C 76 (7) (2016) 359, https://doi.org/10.1140/epjc/s10052-016-4231-5, arXiv:1605.09246.
[69] R. Kleiss, R. Verheyen, J. High Energy Phys. 11 (2017) 182, https://doi.org/10.1007/JHEP11(2017)182, arXiv:1709.04485.
[70] P. Skands, R. Verheyen, Phys. Lett. B 811 (2020) 135878, https://doi.org/10.1016/j.physletb.2020.135878, arXiv:2002.04939.
[71] H. Brooks, P. Skands, Phys. Rev. D 100 (7) (2019) 076006, https://doi.org/10.1103/PhysRevD.100.076006, arXiv:1907.08980.
[72] R. Kleiss, R. Verheyen, Eur. Phys. J. C 80 (10) (2020) 980, https://doi.org/10.1140/epjc/s10052-020-08510-w, arXiv:2002.09248.
[73] S. Catani, B.R. Webber, G. Marchesini, Nucl. Phys. B 349 (1991) 635–654, https://doi.org/10.1016/0550-3213(91)90390-J.

6

Accelerating One-Loop Calculations

Fully-differential NLO calculations define the state-of-the-art for virtually all standard-model and many beyond-the-standard-model processes, not only as stand-alone calculations, but more importantly in the context of NLO matching [338–341] and NLO merging [354–362, 365] schemes, *cf.* sections 3.3.2 and 3.3.3 respectively. Apart from the tree-level matrix elements for the Born and real correction (Born+1-jet), an NLO calculation requires the input of the Born-level one-loop matrix element. These are typically provided by automated tools [131, 133, 205, 206] interfaced to multi-purpose event generators via the Binoth Les Houches accord [214, 215]. With the advent of fully-differential NNLO calculations for many important standard-model processes, one-loop amplitudes are also needed for the real-virtual, *i.e.*, Born+1-jet one-loop, correction. In this context, the fast and numerically stable evaluation in configurations close to singular limits is of utmost importance to deliver reliable results at high statistical significance.

Despite their flexibility, automated one-loop providers employ rather computationally expensive algorithms. While automated approaches are often the only option for calculations in high-multiplicity processes, it is computationally advantageous to exploit analytic results whenever these are available. However, despite analytic results being in principle known for a wide range of processes, it often lacks publicly available code implementations in formats amenable to common event-generation frameworks. Although some analytic one-loop libraries exist, *cf. e.g.* [426, 427], these are either limited to a certain class of processes or not public. A notable exception is the well-established and publicly available MCFM parton-level event generator, which uses analytic matrix elements for all its NLO and NNLO calculations. MCFM offers a large spectrum of processes, but its Fortran matrix-element routines have so-far not generally been made available for external use, especially as most modern tools are written in C++.

Section 6.1 presents an interface to MCFM, giving access to its extensive library of analytic one-loop matrix elements. The interface is written in C++ in a generic fashion so that it is simple to adopt it in most Monte Carlo event generators available today. As a proof of generality, it has been interfaced to the SHERPA and PYTHIA event-generation frameworks. The computational gains to be had with analytic one-loop matrix elements are highlighted in comparisons to three widely-used automated tools both in a stand-alone setup and a typical NLO multi-jet merging setup with SHERPA. The use of analytic one-loop matrix elements provides a so-far little explored but straight-forward avenue to increase the efficiency in Monte Carlo event generation.

6.1 Published Material

Accelerating LHC phenomenology with analytic one-loop amplitudes: A C++ interface to MCFM

John M Campbell, Stefan Höche, and **Christian T Preuss**

Submitted to EPJC
e-Print: arXiv:2107.04472 [hep-ph]

Since the submission of this thesis, a revised version of the paper has been accepted for publication by EPJC, containing an additional section on numerical stability.

Paper begins overleaf.

FERMILAB-PUB-21-290-T, MCNET-21-13

Accelerating LHC phenomenology with analytic one-loop amplitudes: A C++ interface to MCFM

John M Campbell,[1] Stefan Höche,[1] and Christian T Preuss[2]

[1] *Fermi National Accelerator Laboratory, Batavia, IL, 60510, USA*
[2] *School of Physics and Astronomy, Monash University, Wellington Road, Clayton, VIC-3800, Australia*

The evaluation of one-loop matrix elements is one of the main bottlenecks in precision calculations for the high-luminosity phase of the Large Hadron Collider. To alleviate this problem, a new C++ interface to the MCFM parton-level Monte Carlo is introduced, giving access to an extensive library of analytic results for one-loop amplitudes. Timing comparisons are presented for a large set of Standard Model processes. These are relevant for high-statistics event simulation in the context of experimental analyses and precision fixed-order computations.

I. INTRODUCTION

Many measurements at particle colliders can only be made with the help of precise Standard Model predictions, which are typically derived using fixed-order perturbation theory at the next-to-leading order (NLO) or next-to-next-to-leading order (NNLO) in the strong and/or electroweak coupling. Unitarity-based techniques and improvements in tensor reduction during the past two decades have enabled the computation of many new one-loop matrix elements, often using fully numeric techniques [1–19]. The algorithmic appeal and comparable simplicity of the novel approaches has also led to the partial automation of the computation of one-loop matrix elements in arbitrary theories, including effective field theories that encapsulate the phenomenology of a broad range of additions to the Standard Model [20, 21]. With this "NLO revolution" precision phenomenology has entered a new era.

It has become clear, however, that the fully numeric computation of one-loop matrix elements is not without its drawbacks, the most relevant being a relatively large computational complexity. While the best methods exhibit good scaling with the number of final-state particles and are the only means to perform very high multiplicity calculations, it is prudent to resort to known analytic results whenever they are available and computational resources are scarce. The problem has become pressing due to the fact that the computing power on the Worldwide LHC Computing Grid (WLCG) is projected to fall short of the demand by at least a factor two in the high-luminosity phase of the Large Hadron Collider (LHC) [22–25]. Moreover, most techniques for fully differential NNLO calculations rely on the fast and numerically stable evaluation of one-loop results in infrared-singular regions of phase space, further increasing the demand for efficient one-loop computations [26, 27].

In this letter, we report on an extension of the well-known NLO parton-level program MCFM [27–30], which allows the one-loop matrix elements in MCFM to be accessed using the Binoth Les Houches Accord (BLHA) [31, 32] via a direct C++ interface[1]. This is in the same spirit as the BLHA interface to the BLACKHAT library [1], which gives access to analytic matrix elements for V+jet(s), $\gamma\gamma$(+jet) and di-(tri-)jet production. We have constructed the new interface for the most relevant Standard-Model processes available in MCFM, representing a selection of $2 \to n$ processes with $n \leq 4$. As a proof of generality, we have implemented it in the SHERPA [33] and PYTHIA [34] event generation frameworks[2]. We test the newly developed methods in both a stand-alone setup and a typical setup of the SHERPA event generator, and summarize the speed gains in comparison to automated one-loop programs.

II. AVAILABLE PROCESSES

The Standard Model processes currently available through the MCFM one-loop interface are listed in Tab. I, with additional processes available in the Higgs effective theory shown in Tab. II. All processes are implemented in a crossing-invariant fashion. As well as processes available in the most recent version of the MCFM code (v10.0), the interface also allows access to previously unreleased matrix elements for $pp \to \gamma jj$ [38] and di-jet production. Further processes listed in the MCFM manual [39] may be included upon request.

In assembling the interface we have modified the original MCFM routines such that, as far as possible, overhead associated with the calculation of all partonic channels – as required for the normal operation of the MCFM code – is avoided, and only the specific channel that is requested is computed. Additionally, all matrix elements are calculated using the complex-mass scheme [40, 41] and a non-diagonal form of the CKM matrix may be specified in the interface. In general, effects due to loops containing a massive top quark are fully taken into account,

[1] The source code is available at gitlab.com/mcfm-team/releases.
[2] The PYTHIA version has been tested in the context of NLO matrix-element corrections (cf. [35, 36]) in the VINCIA shower [37]. The implementation of NLO MECs in VINCIA and the MCFM interface are planned to be made public in a future PYTHIA 8.3 release.

TABLE I. Processes available in the Standard Model.

Process	Order EW	Order QCD	Reference
$pp \to \ell^+\ell^-$	2	1	–
$pp \to \ell^+\ell^- j$	2	2	[42, 43]
$pp \to \ell^+\ell^- jj$	2	3	[42, 43]
$pp \to \ell^\pm\nu_\ell$	2	1	–
$pp \to \ell^\pm\nu_\ell j$	2	2	[42, 43]
$pp \to \ell^\pm\nu_\ell jj$	2	3	[42, 43]
$pp \to h_0$	1	2	–
$pp \to h_0 j$	1	3	[44]
$pp \to h_0 jj$	1	4	[45, 46]
$pp \to h_0 h_0$	2	2	[47]
$pp \to \ell^+\ell^- h_0$	3	1	–
$pp \to \ell^+\ell^- h_0 j$	3	2	[48]
$pp \to \ell^\pm\nu_\ell h_0$	3	1	–
$pp \to \ell^\pm\nu_\ell h_0 j$	3	2	[48]
$pp \to \gamma j$	1	2	[49, 50]
$pp \to \gamma jj$	1	3	[38]
$pp \to \gamma\gamma$	2	1	–
$gg \to \gamma\gamma$	2	2	[51]
$pp \to \gamma\gamma j$	2	2	[52]
$pp \to \gamma\gamma\gamma$	3	1	[52]
$pp \to \gamma\gamma\gamma\gamma$	4	1	[53]
$pp \to \ell^+\ell^-\gamma$	3	1	[28, 54]
$pp \to \ell^\pm\nu_\ell\gamma$	3	1	[28, 54]
$pp \to \nu_\ell\bar{\nu}_\ell\gamma$	3	1	[28, 54]
$pp \to \ell^+\ell'^-\nu_\ell\bar{\nu}_{\ell'}$	4	1	[28, 54]
$pp \to \ell^+\ell^-\nu_{\ell'}\bar{\nu}_{\ell'}$	4	1	[28, 54]
$pp \to \ell^+\ell^-\ell'^+\ell'^-$	4	1	[28, 54]
$pp \to \ell^+\ell^+\ell^-\ell^-$	4	1	[28, 54]
$pp \to \ell^+\ell^-\ell'^\pm\nu_{\ell'}$	4	1	[28, 54]
$pp \to \ell^\pm\nu_\ell\nu_{\ell'}\bar{\nu}_{\ell'}$	4	1	[28, 54]
$pp \to t\bar{t}$	0	3	[55]
$pp \to jj$	0	3	[56]

TABLE II. Processes available in the Higgs EFT.

Process	Order EW	Order QCD	Reference
$pp \to h_0$	1	2	–
$pp \to h_0 j$	1	3	[57]
$pp \to h_0 jj$	1	4	[58–65]

with the additional requirement that the width of the top quark is set to zero.[3] The intent is that the interface can therefore be used as a direct replacement for a numerical one-loop provider (OLP). We have checked, on a point-by-point basis, that the one-loop matrix elements returned by the interface agree perfectly with those provided by OpenLoops2, Recola2 and MadLoop5. A brief overview of the structure of the interface is given in Appendix B.

III. TIMING BENCHMARKS

To gauge the efficiency gains compared to automated one-loop providers, we compare the evaluation time in Mcfm against OpenLoops2, Recola2, and MadLoop5. The tests are conducted in three stages. First, we test the CPU time needed for the evaluation of loop matrix elements at single phase space points; in a second stage, we test the speedup in the calculation of Born-plus-virtual contributions of NLO calculations using realistic setups; lastly, we compare the CPU time of the different OLPs in a realistic multi-jet merged calculation. In all cases, we estimate the dependence on the computing hardware by running all tests on a total of four different CPUs, namely

- Intel® Xeon® E5-2650 v2 (2.60GHz, 20MB)
- Intel® Xeon® Gold 6150 (2.70GHz, 24.75MB)
- Intel® Xeon® Platinum 8260 (2.40GHz, 35.75MB)
- Intel® Xeon Phi™ 7210 (1.30GHz, 32MB)

For the timing tests at matrix-element level, we use stand-alone interfaces to the respective tools and sample phase space points flatly using the Rambo algorithm [66]. We do not include the time needed for phase-space point generation in our results, and we evaluate a factor 10 more phase-space points in Mcfm in order to obtain more accurate timing measurements at low final-state multiplicity. The main programs and scripts we used for this set of tests are publicly available[1]. The results are collected in Fig. 1, where we show all distinct partonic configurations that contribute to the processes listed in Tabs. I and II. We use the average across the different CPUs as the central value, while the error bars range from the minimal to the maximal value. The interface to Mcfm typically evaluates matrix elements a factor 10–100 faster than the numerical one-loop providers, although for a handful of (low multiplicity) cases this factor can be in the 1,000–10,000 range.

We perform a second set of tests, using the Sherpa event generator [33, 67], its existing OLP interfaces to OpenLoops2 and Recola2 [4] [5] [68], and a dedicated interface to MadLoop5 [6]. With these interfaces we test

[3] An approximate form for top-quark loops is used for the processes $qq' \to \gamma qq'$, $qg \to \gamma qg$, $q\bar{q} \to e^-e^+hg$ and $gg \to gg$, so that strict agreement with other OLPs for these processes requires setting the top mass very large, e.g. to 100 TeV.

[4] At the time of this study, Sherpa provided an interface to Recola's Standard Model implementation only.

[5] For $V+2j$ processes, we use Recola1 due to compatibility issues with Recola2.

[6] We thank Valentin Hirschi for his help in constructing a dedicated MadLoop5 interface to Sherpa. This interface will be described in detail elsewhere.

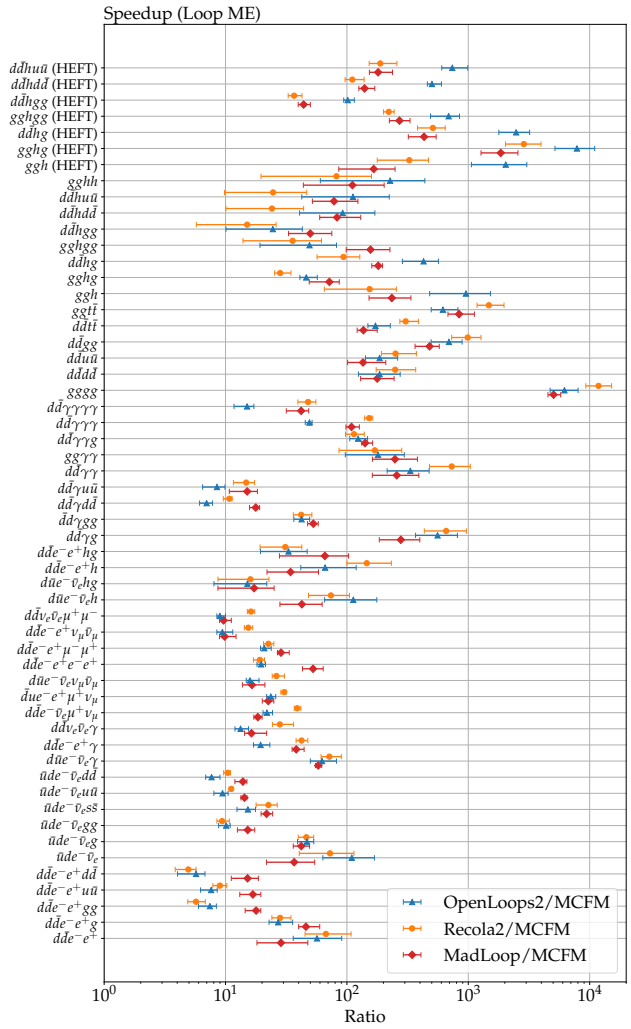


FIG. 1. CPU time ratio of OpenLoops, Recola, and MadLoop5 to Mcfm at the level of loop matrix elements.

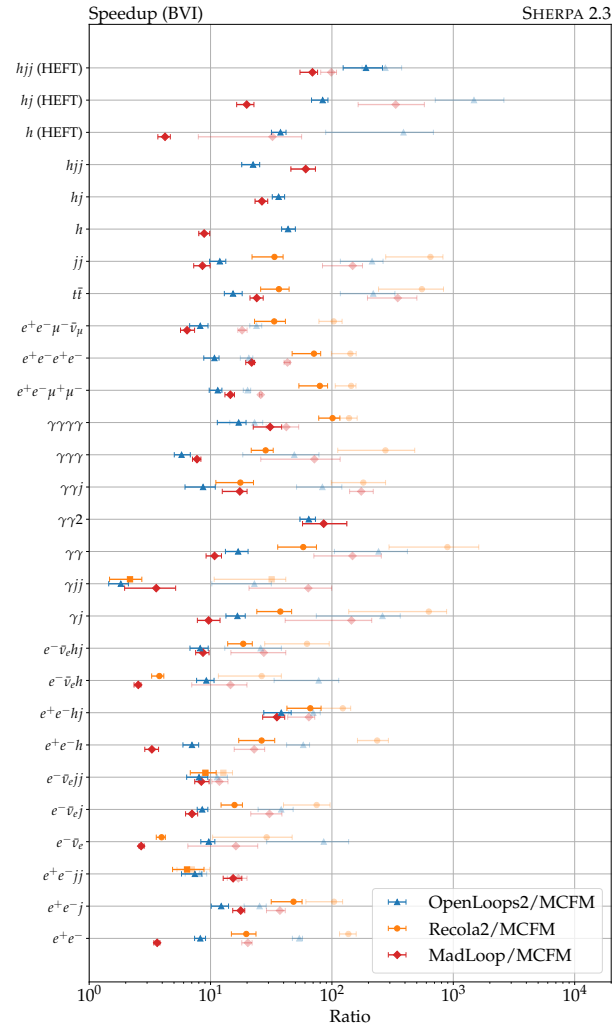


FIG. 2. CPU time ratio of OpenLoops2, Recola2 [5], and MadLoop5 to Mcfm at the level of Born-like contributions to the NLO cross section (BVI).

the speedup in the calculation of the Born-like contributions to a typical NLO computation for the LHC at $\sqrt{s} = 14$ TeV, involving the loop matrix elements in Tabs. I and II. The scale choices and phase-space cuts used in these calculations are listed in App. A. Figure 2 shows the respective timing ratios. It is apparent that the large gains observed in Fig. 1 persist in this setup, because the Born-like contributions to the NLO cross section consist of the Born, integrated subtraction terms, collinear mass factorization counterterms and virtual corrections (BVI), and the timing is dominated by the loop matrix elements if at least one parton is present in the final state at Born level. The usage of Mcfm speeds up the calculation by a large factor compared to the automated OLPs, with the exception of very simple processes, such as $pp \to \ell\bar{\ell}$, $pp \to h$, etc., where the overhead from process management and integration in Sherpa dominates. To assess this overhead we also compute the timing ratios after subtracting the time that the Sherpa computation would take without a loop matrix element. The corresponding results are shown in a lighter shade and confirm that the Sherpa overhead is significant at low multiplicity and becomes irrelevant at higher multiplicity.

In the final set of tests we investigate a typical use case in the context of parton-level event generation for LHC experiments. We use the Sherpa event generator in a multi-jet merging setup for $pp \to W$+jets and $pp \to Z$+jets [69] at $\sqrt{s} = 8$ TeV, with a jet separation cut of $Q_{\text{cut}} = 20$ GeV, and a maximum number of five final state jets at the matrix-element level. Up to two-jet final states are computed at NLO accuracy. In this use case, the gains observed in Figs. 1 and 2 will be greatly diminished, because the timing is dominated by the event generation efficiency for the highest multiplicity tree-level matrix elements [70] and influenced by particle-level event generation as well as the clustering algorithm

TABLE III. CPU time ratios in an NLO multi-jet merged setup using SHERPA.

Merged Process $n \leq 2$ @ NLO $n \leq 5$ @ LO	SHERPA + OPENLOOPS2/MCFM	SHERPA + MADLOOP5/MCFM
$pp \to Z + nj$	$1.83^{+0.20}_{-0.12}$	$3.01^{+0.26}_{-0.18}$
$pp \to W^+ + nj$	$1.34^{+0.06}_{-0.07}$	$1.36^{+0.03}_{-0.03}$
$pp \to W^- + nj$	$1.38^{+0.06}_{-0.04}$	$1.38^{+0.07}_{-0.11}$

needed for multi-jet merging[7]. We make use of the efficiency improvements described in Ref. [73], in particular neglecting color and spin correlations in the S-MC@NLO matching procedure [74]. We do not include underlying event simulation or hadronization. The results in Tab. III still show a fairly substantial speedup when using MCFM. We point out that a higher gain could be achieved by also making use of MCFM's implementation of analytic matrix elements for real-emission corrections and Catani-Seymour dipole terms.

We close this section with a direct comparison of the CPU time needed for the calculation of Drell-Yan processes with one and two jets using SHERPA and MCFM, up to a target precision on the integration of 0.1% (one jet) or 0.3% (two jets). The center-of-mass energy is $\sqrt{s} = 14$ TeV, and the scale choices and cuts are listed in App. A. The results are shown in Tab. IV. As might be expected when comparing a dedicated parton-level code with a general-purpose particle-level generator, MCFM is substantially faster than SHERPA for the evaluation of all contributions to the NLO calculation. These results indicate a few avenues for further improvements of general-purpose event generators. With the efficient evaluation of virtual contributions in hand, attention should now turn to the calculation of real-radiation configurations – that represent the bottleneck for both SHERPA and MCFM. In the simplest cases with up to 5 partons, the real radiation and dipole counterterms could be evaluated using analytic rather than numerical matrix elements, by a suitable extension of the interface we have presented here. In addition, the form of the phase-space generation may be improved for Born-like phase-space integrals. Table IV lists the number of phase-space points before cuts that are required to achieve the target accuracy. We find that MCFM uses fewer than half of the points needed by SHERPA in the Born-like phase-space integrals, while SHERPA uses fewer points than MCFM in the real-emission type integrals but at a much higher computational cost. This confirms that SHERPA's event generation is indeed impaired by the slow evaluation of real-emission type matrix elements, and by the factorial scaling of the diagram-based phase-space integration technique [75, 76] used in its calculations[8].

[7] In this study we do not address the question of additional timing overhead due to NLO electroweak corrections or PDF reweighting [71], which could both be relevant in practice. It has recently been shown that in good implementations of the reweighting and EW correction algorithm, the additional overhead will not be sizable [72].

TABLE IV. Comparison of integration times using SHERPA and MCFM.

Process MC accuracy		SHERPA time / #pts	MCFM time / #pts
$pp \to Zj$	Born-like	76.8 m / 11.3M	13.6 m /4.5M
0.1%	real-like	38 h / 33.1M	51.5 m / 22.5M
$pp \to Zjj$	Born-like	96.0 h / 22.4M	19.6 h / 4.5M
0.3%	real-like	830.4 h / 58.7M	62.9 h / 83.8M
$pp \to W^+j$	Born-like	40.5 m / 12.8M	7.37 m / 4.5M
0.1%	real-like	16.9 h / 38.3M	59.4 m / 36.0M
$pp \to W^+jj$	Born-like	14.1 h / 20.3M	9.32 h / 7.2M
0.3%	real-like	222.1 h / 38.9M	54.4 h / 119.8M
$pp \to W^-j$	Born-like	34.1 m / 11.0M	7.46 m / 4.5M
0.1%	real-like	15.9 h / 40.5M	47.2 m / 28.1M
$pp \to W^-jj$	Born-like	12.8 h / 20.0M	7.34 h / 5.6M
0.3%	real-like	281.1 h / 52.0M	38.8 h / 83.8M

IV. CONCLUSIONS

We have presented a novel C++ interface to the well-known MCFM parton-level Monte Carlo generator, giving access to its extensive library of analytical one-loop amplitudes. The interface is generic and not tied to any specific Monte Carlo event generation tool. As a proof of its generality, we have implemented the interface in both, the SHERPA and PYTHIA event generators. The SHERPA interface will become public with version 3.0.0, and the PYTHIA interface is foreseen to become public in a future release of the 8.3 series[2]. It should be straightforward to adapt our code to the needs of other event generators.

We expect the interface to be valuable in two respects. First, for many of the processes considered here the speedup over other OLPs is substantial; accessing these matrix elements via this interface rather than an automated tool will therefore provide an immediate acceleration of event generation for many processes of high phenomenological interest. Second, the speed comparisons presented here highlight processes that are particularly computationally intensive for automated tools. Further improvements to the efficiency of these codes may be possible, with potential gains across a wider range of processes.

[8] We do not make use of SHERPA's recursive phase-space generator [77], because it is available for color-sampled matrix element evaluation only. Color sampling would further reduce the efficiency of the integration, because the processes at hand involve a relatively small number of QCD partons.

The structure of the interface allows for simple extensions. Further one-loop matrix elements in MCFM, implemented either currently or in the future, may become accessible in a straightforward manner. In the same spirit, the interface could also be extended to provide tree-level or two-loop matrix elements included in MCFM as the need arises. Further extensions to the interface, for instance to provide finer control over the one-loop matrix elements via the selection of helicities or color configurations, would also be possible.

Given that we have interfaced three popular automated OLPs within the generator-agnostic structure of the new MCFM interface, it is natural to envision the future development of a hybrid program that makes use of the fastest matrix element library for each process. Thinking further ahead, it may be worthwhile to reconsider a streamlined event generation framework, combining different (dedicated) parton-level and particle-level tools. This idea has been pursued with THEPEG [78], but so far rarely deployed. Apart from obvious efficiency improvements through the use of dedicated tools for different applications, such a framework enables previously unavailable methods for systematics studies. In view of both the faster integration in MCFM over SHERPA and the magnitude of uncertainties pertaining to theoretical modeling of collider observables, this is becoming an increasingly important avenue for future work.

We want to close by highlighting that only a relatively small number of analytical amplitudes has to be known in order to cover a wide range of physical processes. When judiciously assembled, many parts of the calculations can be recycled in a process-independent way, with only charge and coupling factors being process-specific. Compared to other efforts to increase the efficiency of event generators, swapping automated for analytical matrix elements is straightforward and simple. Analytical matrix element libraries provide a so-far little explored path towards higher-efficiency event generation for the (high-luminosity) LHC and future colliders.

ACKNOWLEDGMENTS

We are grateful to Frank Krauss and Marek Schönherr for many fruitful discussions. CTP thanks Peter Skands for support. CTP is supported by the Monash Graduate Scholarship, the Monash International Postgraduate Research Scholarship, and the J. L. William Scholarship. This research was supported by the Fermi National Accelerator Laboratory (Fermilab), a U.S. Department of Energy, Office of Science, HEP User Facility. Fermilab is managed by Fermi Research Alliance, LLC (FRA), acting under Contract No. DE–AC02–07CH11359. We further acknowledge support from the Monash eResearch Centre and eSolutions-Research Support Services through the MonARCH HPC Cluster. This work was also supported in part by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 722104 – MCnetITN3.

Appendix A: Parameters and cuts for timing comparisons

In order to perform the timing comparisons shown in Fig. 2 and Table IV. we employ the following scale choices and phase-space cuts:

- $\Delta R_{\ell,\gamma} > 0.4$
- $\Delta R_{\gamma,\gamma} > 0.4$
- $p_{\mathrm{T},\gamma} > 30$ GeV
- $p_{\mathrm{T},j} > 30$ GeV
- 66 GeV $< m_{\ell\bar{\ell}} <$ 116 GeV

We reconstruct jets using the anti-k_{T} algorithm [79] in the implementation of FastJet [80] with an R parameter of 0.4. For the di-jet process we require $p_{\mathrm{T},j}$ >80 GeV. Photons are isolated from QCD activity based on Ref. [81] with δ_0=0.4, n=2 and ϵ_γ=2.5%

Appendix B: Structure of the interface

The MCFM C++ interface is constructed as a C++ class

```
CXX_Interface mcfm;
```

included in the header:

```
#include "MCFM/CXX_Interface.h"
```

It must be initialized on a `std::map` of `std::string`s, containing all (standard-model) parameters:

```
bool CXX_Interface::Initialize(
        std::map<std::string,std::string>& parameters);
```

Prior to use, each process has to be initialized in the interface:

```
int CXX_Interface::InitializeProcess(const Process_Info &pi);
```

which takes a `Process_Info` object as input, which in turn contains the defining parameters of a given process, i.e. the PDG IDs, number of incoming particles, and QCD and EW coupling orders:

```
Process_Info(const std::vector<int> &ids, const int nin,
        const int oqcd, const int oew);
```

Phase space points are defined using the `FourVec` struct, which represents four-vectors in the ordering (E, p_x, p_y, p_z).

```
FourVec(double e, double px, double py, double pz);
```

Given a list of four-vectors in this format, one-loop matrix elements can be calculated either using the process ID returned by the `InitializeProcess` method

```
void CXX_Interface::Calc(int procID,
        const std::vector<FourVec> &p, int oqcd);
```

or using a `Process_Info` struct:

```
void CXX_Interface::Calc(const Process_Info &pi,
        const std::vector<FourVec> &p,int oqcd);
```

In the same way, the result of this calculation can be accessed either via the process ID

```
const std::vector<double> &CXX_Interface::GetResult(int procID);
```

or using the `Process_Info` struct:

```
const std::vector<double> &CXX_Interface::GetResult(const Process_Info &pi)
```

The result is returned as a list of Laurent series coefficients in the format $(\mathcal{O}(\varepsilon^0), \mathcal{O}(\varepsilon^{-1}), \mathcal{O}(\varepsilon^{-2}), \text{Born})$. However, by default only the $\mathcal{O}(\varepsilon^0)$ coefficient, i.e. the finite part, is returned. The calculation of the pole terms and the Born can be enabled by setting the following switch to 1:

```
void CXX_Interface::SetPoleCheck(int check);
```

An example code showing the basic usage of the interface as well as a function filling the complete list of parameters with default values is publicly available[1].

[1] C. F. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, H. Ita, D. A. Kosower, and D. Maître, Phys. Rev. D **78**, 036003 (2008), arXiv:0803.4180 [hep-ph].

[2] C. F. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, T. Gleisberg, H. Ita, D. A. Kosower, and D. Maître, Phys. Rev. D **80**, 074036 (2009), arXiv:0907.1984 [hep-ph].

[3] C. F. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, T. Gleisberg, H. Ita, D. A. Kosower, and D. Maître, Phys. Rev. D **82**, 074002 (2010), arXiv:1004.1659 [hep-ph].

[4] C. F. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, T. Gleisberg, H. Ita, D. A. Kosower, and D. Maître, Phys. Rev. Lett. **106**, 092001 (2011), arXiv:1009.2338 [hep-ph].

[5] H. Ita, Z. Bern, L. J. Dixon, F. Febres Cordero, D. A. Kosower, and D. Maître, Phys. Rev. D **85**, 031501 (2012), arXiv:1108.2229 [hep-ph].

[6] Z. Bern, L. J. Dixon, F. Febres Cordero, S. Höche, H. Ita, D. A. Kosower, D. Maître, and K. J. Ozeren, Phys. Rev. D **88**, 014025 (2013), arXiv:1304.1253 [hep-ph].

[7] V. Hirschi, R. Frederix, S. Frixione, M. V. Garzelli, F. Maltoni, and R. Pittau, JHEP **05**, 044 (2011), arXiv:1103.0621 [hep-ph].

[8] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, JHEP **07**, 079 (2014), arXiv:1405.0301 [hep-ph].

[9] F. Cascioli, P. Maierhöfer, and S. Pozzorini, Phys. Rev. Lett. **108**, 111601 (2012), arXiv:1111.5206 [hep-ph].

[10] F. Buccioni, S. Pozzorini, and M. Zoller, Eur. Phys. J. C **78**, 70 (2018), arXiv:1710.11452 [hep-ph].

[11] F. Buccioni, J.-N. Lang, J. M. Lindert, P. Maierhöfer, S. Pozzorini, H. Zhang, and M. F. Zoller, Eur. Phys. J. C **79**, 866 (2019), arXiv:1907.13071 [hep-ph].

[12] S. Badger, B. Biedermann, and P. Uwer, Comput. Phys. Commun. **182**, 1674 (2011), arXiv:1011.2900 [hep-ph].

[13] S. Badger, B. Biedermann, P. Uwer, and V. Yundin, Comput. Phys. Commun. **184**, 1981 (2013), arXiv:1209.0100 [hep-ph].

[14] G. Cullen, N. Greiner, G. Heinrich, G. Luisoni, P. Mastrolia, G. Ossola, T. Reiter, and F. Tramontano, Eur. Phys. J. C **72**, 1889 (2012), arXiv:1111.2034 [hep-ph].

[15] G. Cullen *et al.*, Eur. Phys. J. C **74**, 3001 (2014), arXiv:1404.7096 [hep-ph].

[16] S. Actis, A. Denner, L. Hofer, A. Scharf, and S. Uccirati, JHEP **04**, 037 (2013), arXiv:1211.6316 [hep-ph].

[17] S. Actis, A. Denner, L. Hofer, J.-N. Lang, A. Scharf, and S. Uccirati, Comput. Phys. Commun. **214**, 140 (2017), arXiv:1605.01090 [hep-ph].

[18] A. Denner, J.-N. Lang, and S. Uccirati, JHEP **07**, 087 (2017), arXiv:1705.06053 [hep-ph].

[19] A. Denner, J.-N. Lang, and S. Uccirati, Comput. Phys. Commun. **224**, 346 (2018), arXiv:1711.07388 [hep-ph].

[20] C. Degrande, Comput. Phys. Commun. **197**, 239 (2015), arXiv:1406.3030 [hep-ph].

[21] C. Degrande, G. Durieux, F. Maltoni, K. Mimasu, E. Vryonidou, and C. Zhang, (2020), arXiv:2008.11743

[hep-ph].
[22] J. Albrecht *et al.* (HEP Software Foundation), Comput. Softw. Big Sci. **3**, 7 (2019), arXiv:1712.06982 [physics.comp-ph].
[23] S. Alioli *et al.*, (2019), arXiv:1902.01674 [hep-ph].
[24] S. Amoroso *et al.* (HSF Physics Event Generator WG), Comput. Softw. Big Sci. **5**, 12 (2021), arXiv:2004.13687 [hep-ph].
[25] T. Aarrestad *et al.* (HEP Software Foundation), in *2022 Snowmass Summer Study*, edited by P. Canal *et al.* (2020) arXiv:2008.13636 [physics.comp-ph].
[26] R. Boughezal, J. M. Campbell, R. K. Ellis, C. Focke, W. Giele, X. Liu, F. Petriello, and C. Williams, Eur. Phys. J. C **77**, 7 (2017), arXiv:1605.08011 [hep-ph].
[27] J. Campbell and T. Neumann, JHEP **12**, 034 (2019), arXiv:1909.09117 [hep-ph].
[28] J. M. Campbell and R. K. Ellis, Phys. Rev. D **60**, 113006 (1999), arXiv:hep-ph/9905386.
[29] J. M. Campbell, R. K. Ellis, and C. Williams, JHEP **07**, 018 (2011), arXiv:1105.0020 [hep-ph].
[30] J. M. Campbell, R. K. Ellis, and W. T. Giele, Eur. Phys. J. C **75**, 246 (2015), arXiv:1503.06182 [physics.comp-ph].
[31] T. Binoth *et al.*, Comput. Phys. Commun. **181**, 1612 (2010), arXiv:1001.1307 [hep-ph].
[32] S. Alioli *et al.*, Comput. Phys. Commun. **185**, 560 (2014), arXiv:1308.3462 [hep-ph].
[33] E. Bothmann *et al.* (Sherpa), SciPost Phys. **7**, 034 (2019), arXiv:1905.09127 [hep-ph].
[34] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, Comput. Phys. Commun. **191**, 159 (2015), arXiv:1410.3012 [hep-ph].
[35] L. Hartgring, E. Laenen, and P. Skands, JHEP **10**, 127 (2013), arXiv:1303.4974 [hep-ph].
[36] N. Baberuxki, C. T. Preuss, D. Reichelt, and S. Schumann, JHEP **04**, 112 (2020), arXiv:1912.09396 [hep-ph].
[37] H. Brooks, C. T. Preuss, and P. Skands, JHEP **07**, 032 (2020), arXiv:2003.00702 [hep-ph].
[38] J. M. Campbell, R. K. Ellis, and C. Williams, Phys. Rev. Lett. **118**, 222001 (2017), [Erratum: Phys.Rev.Lett. 124, 259901 (2020)], arXiv:1612.04333 [hep-ph].
[39] J. Campbell, K. Ellis, W. Giele, T. Neumann, and C. Williams, https://mcfm.fnal.gov/.
[40] A. Denner, S. Dittmaier, M. Roth, and D. Wackeroth, Nucl. Phys. B **560**, 33 (1999), arXiv:hep-ph/9904472.
[41] A. Denner, S. Dittmaier, M. Roth, and L. H. Wieders, Nucl. Phys. B **724**, 247 (2005), [Erratum: Nucl.Phys.B 854, 504–507 (2012)], arXiv:hep-ph/0505042.
[42] Z. Bern, L. J. Dixon, and D. A. Kosower, Nucl. Phys. B **513**, 3 (1998), arXiv:hep-ph/9708239.
[43] J. M. Campbell and R. K. Ellis, JHEP **01**, 020 (2017), arXiv:1610.02189 [hep-ph].
[44] R. K. Ellis, I. Hinchliffe, M. Soldate, and J. J. van der Bij, Nucl. Phys. B **297**, 221 (1988).
[45] R. K. Ellis and S. Seth, JHEP **11**, 006 (2018), arXiv:1808.09292 [hep-ph].
[46] L. Budge, J. M. Campbell, G. De Laurentis, R. K. Ellis, and S. Seth, JHEP **05**, 079 (2020), arXiv:2002.04018 [hep-ph].
[47] E. Glover and J. J. van der Bij, Nucl. Phys. B **309**, 282 (1988).
[48] J. M. Campbell, R. K. Ellis, and C. Williams, JHEP **06**, 179 (2016), arXiv:1601.00658 [hep-ph].
[49] R. K. Ellis, D. A. Ross, and A. E. Terrano, Nucl. Phys. B **178**, 421 (1981).
[50] P. Aurenche, R. Baier, A. Douiri, M. Fontannaz, and D. Schiff, Nucl. Phys. B **286**, 553 (1987).
[51] J. M. Campbell, R. K. Ellis, Y. Li, and C. Williams, JHEP **07**, 148 (2016), arXiv:1603.02663 [hep-ph].
[52] J. M. Campbell and C. Williams, Phys. Rev. D **89**, 113001 (2014), arXiv:1403.2641 [hep-ph].
[53] T. Dennen and C. Williams, Phys. Rev. D **91**, 054012 (2015), arXiv:1411.3237 [hep-ph].
[54] L. J. Dixon, Z. Kunszt, and A. Signer, Nucl. Phys. B **531**, 3 (1998), arXiv:hep-ph/9803250.
[55] P. Nason, S. Dawson, and R. K. Ellis, Nucl. Phys. B **327**, 49 (1989), [Erratum: Nucl.Phys.B 335, 260–260 (1990)].
[56] R. K. Ellis and J. C. Sexton, Nucl. Phys. B **269**, 445 (1986).
[57] C. R. Schmidt, Phys. Lett. B **413**, 391 (1997), arXiv:hep-ph/9707448.
[58] L. J. Dixon, E. Glover, and V. V. Khoze, JHEP **12**, 015 (2004), arXiv:hep-th/0411092.
[59] R. K. Ellis, W. T. Giele, and G. Zanderighi, Phys. Rev. D **72**, 054018 (2005), [Erratum: Phys.Rev.D 74, 079902 (2006)], arXiv:hep-ph/0506196.
[60] S. D. Badger and E. Glover, Nucl. Phys. B Proc. Suppl. **160**, 71 (2006), arXiv:hep-ph/0607139.
[61] S. D. Badger, E. Glover, and K. Risager, JHEP **07**, 066 (2007), arXiv:0704.3914 [hep-ph].
[62] E. Glover, P. Mastrolia, and C. Williams, JHEP **08**, 017 (2008), arXiv:0804.4149 [hep-ph].
[63] S. Badger, E. Glover, P. Mastrolia, and C. Williams, JHEP **01**, 036 (2010), arXiv:0909.4475 [hep-ph].
[64] L. J. Dixon and Y. Sofianatos, JHEP **08**, 058 (2009), arXiv:0906.0008 [hep-ph].
[65] S. Badger, J. M. Campbell, R. K. Ellis, and C. Williams, JHEP **12**, 035 (2009), arXiv:0910.4481 [hep-ph].
[66] R. Kleiss, W. J. Stirling, and S. D. Ellis, Comput. Phys. Commun. **40**, 359 (1986).
[67] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, JHEP **02**, 007 (2009), arXiv:0811.4622 [hep-ph].
[68] B. Biedermann, S. Bräuer, A. Denner, M. Pellen, S. Schumann, and J. M. Thompson, Eur. Phys. J. C **77**, 492 (2017), arXiv:1704.05783 [hep-ph].
[69] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, JHEP **04**, 027 (2013), arXiv:1207.5030 [hep-ph].
[70] S. Höche, S. Prestel, and H. Schulz, Phys. Rev. D **100**, 014024 (2019), arXiv:1905.05120 [hep-ph].
[71] E. Bothmann, M. Schönherr, and S. Schumann, Eur. Phys. J. C **76**, 590 (2016), arXiv:1606.08753 [hep-ph].
[72] E. Bothmann, S. Höche, and M. Schönherr, (2021).
[73] K. Danziger, (2020), https://cds.cern.ch/record/2715727.
[74] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, JHEP **09**, 049 (2012), arXiv:1111.1220 [hep-ph].
[75] E. Byckling and K. Kajantie, Nucl. Phys. **B9**, 568 (1969).
[76] E. Byckling and K. Kajantie, Phys.Rev. **187**, 2008 (1969).
[77] T. Gleisberg and S. Höche, JHEP **12**, 039 (2008), arXiv:0808.3674 [hep-ph].
[78] L. Lönnblad, Nucl. Instrum. Meth. A **559**, 246 (2006).
[79] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **04**, 063 (2008), arXiv:0802.1189 [hep-ph].
[80] M. Cacciari, G. P. Salam, and G. Soyez, Eur. Phys. J. C **72**, 1896 (2012), arXiv:1111.6097 [hep-ph].
[81] S. Frixione, Phys. Lett. B **429**, 369 (1998), arXiv:hep-ph/9801442.

7

Towards NNLO Matching with Sector Showers

With the advancement of the general and automated fully-differential NLO matching methods MC@NLO [115, 338, 342] and POWHEG [229, 339–341, 428], NLO-plus-parton-shower calculations have become state-of-the-art for collider-physics simulations. As fully-differential NNLO calculations are becoming available for a multitude of processes, it is prudent to consider the combination of these with parton-shower simulations. Via the matching to parton showers, higher-order calculations can be embedded in multi-pupose event-generation frameworks, which offer the modelling of full particle-level collider events, including multi-parton interactions (*cf.* section 3.1.5) and hadronisation effects (*cf.* section 3.4). The detailed and precise simulation of collider events is especially relevant for so-called background processes, in order to facilitate the discrimination of known standard-model and new-physics phenomena.

However, fully-differential matching methods are so far not available beyond NLO, because existing parton showers do not capture the correct singularity structure of corrections at higher orders, despite the fact that significant progress has been made on including higher-order corrections in parton-shower algorithms [275, 276, 312, 313]. Nevertheless, techniques have been developed to combine NNLO calculations with parton-shower simulations [362, 366, 367, 369, 372–378, 429, 430]. However, these do not offer the same level of control and flexibility as the methods available at NLO and as would be desirable for high-precision predictions.

In the manuscript included in section 7.1, a first step towards fully-differential matching at NNLO is presented, restricting the discussion to processes with colour-singlet initial-state particles and two final-state jets. The method extends the idea developed at NLO in [345] and may therefore be viewed as an extension of the POWHEG scheme. Specifically, it does not require the shower to provide the fixed-order subtraction terms (as needed in MC@NLO-type approaches) but uses the shower as an efficient Sudakov-weighted phase-space generator. The implementation in VINCIA presented in section 7.1 is based on the NNLO antenna-subtraction method and the sector-shower framework, with the latter amended by real-virtual and double-real corrections. As such, it combines and generalises multiple developments undertaken within the VINCIA framework over the past decade: iterated tree-level matrix-element corrections [277,336], NLO matrix-element corrections [322], and second-order shower branchings [275].

7.1 Published Material

Towards NNLO+PS Matching with Sector Showers

John M Campbell, Stefan Höche, Hai Tao Li, **Christian T Preuss**, and Peter Skands

Submitted to Physics Letters B
e-Print: arXiv:2108.07133 [hep-ph]

Paper begins overleaf.

Towards NNLO+PS Matching with Sector Showers

John M Campbell[a], Stefan Höche[a], Hai Tao Li[b,c], Christian T Preuss[d], Peter Skands[d]

[a]*Fermi National Accelerator Laboratory, Batavia, IL, 60510, USA*
[b]*HEP Division, Argonne National Laboratory, Argonne, Illinois 60439, USA*
[c]*Department of Physics & Astronomy, Northwestern University, Evanston, IL 60208, USA*
[d]*School of Physics and Astronomy, Monash University, Wellington Road, Clayton, VIC-3800, Australia*

Abstract

We outline a new technique for the fully-differential matching of final-state parton showers to NNLO calculations, focussing here on the simplest case of leptonic collisions with two final-state jets. The strategy is facilitated by working in the antenna formalism, making use of NNLO antenna subtraction on the fixed-order side and the sector-antenna framework on the shower side. As long as the combined real-virtual and double-real corrections do not overcompensate the real-emission term in the three-jet region, negative weights can be eliminated from the matching scheme. We describe the implementation of all necessary components in the VINCIA antenna shower in PYTHIA 8.3.

Keywords: NNLO matching, parton showers, antenna subtraction, antenna showers

1. Introduction

To date it is possible to perform most collider-physics studies with fully-differential NLO+PS matching thanks to two general, well-developed, and widely applied procedures: MC@NLO [1] and POWHEG [2]. By fully differential matching, we understand that the matching is done point by point in both the Born- and real-emission phase spaces, with a parton shower that reflects the correct singular structure of the fixed-order calculation. In this sense, fully-differential matching requires the fixed-order expansion of the shower to develop the same singularities as the fixed-order calculation up to the matched order. At NLO, this is achieved by parton showers that exponentiate terms reducing to the universal DGLAP kernels in any collinear limit and the eikonal factor in soft limits, with the colour dependence in the soft limit requiring special attention [3]. As of today, no fully-differential matching method obeying these criteria is available at NNLO, although significant progress on including higher-order corrections to parton showers has been made [4–7].

Existing NNLO+PS matching methods either extend existing merging schemes or utilise analytical resummation for the transition between the fixed-order and parton-shower realms. Examples of the first kind are UN^{2}LOPS [8], which extends the UNLOPS [9] scheme to the second order, and MiNNLO$_{\rm PS}$ [10, 11] as well as other extensions [12–18] of the MiNLO technique [19]. The UN^{2}LOPS scheme has recently been generalised to processes with an additional jet in the context of an UNLOPS-based N^{3}LO+PS matching strategy [20]. The MiNLO-based schemes may be seen as a hybrid approach, since they use a combination of analytical and numerical resummation. A noteworthy example of a scheme employing the latter approach is implemented in the GENEVA framework [21, 22]. While all of these have enabled impressive phenomenological studies [8, 23–32] and provide pathways to matching precision calculations to event generators, they do not provide the same level of theoretical control as the fully-differential matching methods that are available at NLO.

In this letter, we present for the first time a fully-differential NNLO+PS matching scheme for final-state parton showers, restricting ourselves to the case of two coloured final-state particles. The new method combines NNLO antenna subtraction with the sector-antenna shower in VINCIA [33], suitably extended to include real-virtual and double-real corrections. A key aspect of the new technique is that the parton shower is employed only as an efficient Sudakov-weighted phase-space generator. It does not define the infrared sub-

Email addresses: johnmc@fnal.gov (John M Campbell), shoeche@fnal.gov (Stefan Höche), haitao.li@northwestern.edu (Hai Tao Li), christian.preuss@monash.edu (Christian T Preuss), peter.skands@monash.edu (Peter Skands)

traction terms that are key to MC@NLO type matching strategies.

The letter is structured as follows. We review the matching method at NLO in section 2 before extending it to the NNLO in section 3, retaining a rather general notation. A numerical implementation in the VINCIA sector-antenna shower in PYTHIA 8.3 [34] is described in section 4, featuring a more detailed description of the matching scheme. We conclude in section 5 and provide an outlook on applications beyond the simple cases considered here.

2. NLO Matching Strategy

Our matching strategy generalises the technique first developed in [35], which nowadays is referred to as the POWHEG scheme [2, 36, 37]. To start with, it is thus useful to recap the NLO matching strategy, before moving on to the new NNLO technique.

At NLO, the expected value of an infrared-safe observable O defined on a two-particle final state process with a colourless initial state is given by

$$\langle O\rangle_{\mathrm{NLO}} = \int \mathrm{d}\Phi_2 \left[\mathrm{B}(\Phi_2) + \mathrm{V}(\Phi_2) + \mathrm{I}_{\mathrm{S}}^{\mathrm{NLO}}(\Phi_2)\right] O(\Phi_2) + \int \mathrm{d}\Phi_3 \left[\mathrm{R}(\Phi_3)O(\Phi_3) - \mathrm{S}^{\mathrm{NLO}}(\Phi_3)O(\Phi_2(\Phi_3))\right], \quad (1)$$

where B and V denote the Born cross section and virtual correction, differential in the two-particle phase space Φ_2. Similarly, R denotes the real-radiation cross section differential in the three-particle phase space Φ_3, and $\mathrm{S}^{\mathrm{NLO}}$ denotes the differential NLO subtraction term in the antenna subtraction method, with its integral over the antenna phase space given by $\mathrm{I}_{\mathrm{S}}^{\mathrm{NLO}}$.[1] In order to achieve a Born-local cancellation of the subtraction term upon integration over the real-emission phase space, the observable acting on $\mathrm{S}^{\mathrm{NLO}}$ must be evaluated at the reduced phase-space point $\Phi_2(\Phi_3)$, where the precise mapping from the three-parton to the two-parton state depends on the subtraction scheme. We can invert this mapping and factorise the phase space into the 2-particle (Born) phase space Φ_2, and the one-particle radiation phase space Φ_{+1},

$$\mathrm{d}\Phi_3 = \mathrm{d}\Phi_2 \times \mathrm{d}\Phi_{+1}\,. \quad (2)$$

By defining a Born-local NLO weight,

$$k_{\mathrm{NLO}}(\Phi_2) := 1 + \frac{\mathrm{V}(\Phi_2)}{\mathrm{B}(\Phi_2)} + \frac{\mathrm{I}_{\mathrm{S}}^{\mathrm{NLO}}(\Phi_2)}{\mathrm{B}(\Phi_2)} + \int \mathrm{d}\Phi_{+1} \left[\frac{\mathrm{R}(\Phi_2,\Phi_{+1})}{\mathrm{B}(\Phi_2)} - \frac{\mathrm{S}^{\mathrm{NLO}}(\Phi_2,\Phi_{+1})}{\mathrm{B}(\Phi_2)}\right] \quad (3)$$

eq. (1) can be rewritten as

$$\langle O\rangle_{\mathrm{NLO}} = \int \mathrm{d}\Phi_2\, \mathrm{B}(\Phi_2) \left[k_{\mathrm{NLO}}(\Phi_2)O(\Phi_2) + \int \mathrm{d}\Phi_{+1}\, \frac{\mathrm{R}(\Phi_2,\Phi_{+1})}{\mathrm{B}(\Phi_2)} \left(O(\Phi_2,\Phi_{+1}) - O(\Phi_2)\right)\right]. \quad (4)$$

The parton-shower evolution, on the other hand, is described by a generating functional, the shower operator, recursively defined for an infrared-safe observable O by

$$\mathcal{S}_n(t,O) = \Delta_n(t,t_{\mathrm{c}})O(\Phi_n) + \int_{t_{\mathrm{c}}}^{t} \mathrm{d}\Phi_{+1}\, \mathrm{A}_n^{(0)}(\Phi_{+1})\Delta_n(t,t')\mathcal{S}_{n+1}(t',O)\,, \quad (5)$$

where $\mathrm{A}_n^{(0)}(\Phi_{+1})$ is the sum of all leading-order antenna functions[2] competing for the next branching $IK \mapsto ijk$ off the n-parton configuration,

$$\int_{t_{\mathrm{c}}}^{t} \mathrm{d}\Phi_{+1}\, \mathrm{A}^{(0)}_{n\mapsto n+1}(\Phi_{+1}) := \int_{t_{\mathrm{c}}}^{t} \sum_{j\in\{n\mapsto n+1\}} A^{(0)}_{j/IK}\left(\Phi^j_{+1}\right)\, \mathrm{d}\Phi^j_{+1}\,, = \sum_{j\in\{n\mapsto n+1\}} \int_{t_{\mathrm{c}}}^{t} \frac{\alpha_{\mathrm{s}}(t)}{4\pi} C_{j/IK}\bar{A}^{(0)}_{j/IK}(t,\zeta,\phi)\,\mathrm{d}t\,\mathrm{d}\zeta\,\frac{\mathrm{d}\phi}{2\pi} \quad (6)$$

with the sum and shower variables left implicit in our notation. Note that when working in the sector antenna framework [33, 40–43], Eq. (6) implicitly defines a partitioning of the real-emission term along with the associated subtractions in Eq. (3). This is crucial to avoid double-counting of radiative corrections generated by the parton shower. The associated Sudakov factor is given by

$$\Delta_n^{\mathrm{LO}}(t_0,t) = \exp\left\{-\int_t^{t_0} \mathrm{d}\Phi_{+1}\, \mathrm{A}^{(0)}_{n\mapsto n+1}(\Phi_{+1})\right\}. \quad (7)$$

[1] Other well-established NLO subtraction schemes such as FKS [38] or dipole subtraction [39] may equally well be employed here.

[2] We refer to NLO antenna subtraction terms as LO antenna functions.

Taking only the first shower emission into account, the expected value of the observable O at LO is given by

$$\langle O\rangle_{\text{LO+PS}} = \int \mathrm{d}\Phi_2\, \mathrm{B}(\Phi_2)\Big[\Delta(t_0,t_c)O(\Phi_2) + \int \mathrm{d}\Phi_{+1}\, \mathrm{A}^{(0)}_{2\mapsto 3}(\Phi_{+1})\Delta(t_0,t)O(\Phi_2,\Phi_{+1})\Big].$$

This implies that upon the replacement

$$\begin{aligned}\mathrm{B}(\Phi_2) &\to k_{\text{NLO}}(\Phi_2)\mathrm{B}(\Phi_2)\\ \mathrm{A}^{(0)}_{2\mapsto 3} &\to w^{\text{LO}}_{2\mapsto 3}(\Phi_2,\Phi_{+1})\mathrm{A}^{(0)}_{2\mapsto 3}\end{aligned} \tag{8}$$

where we have defined the $2 \mapsto 3$ LO matrix-element correction factor,

$$w^{\text{LO}}_{2\mapsto 3}(\Phi_2,\Phi_{+1}) = \frac{\mathrm{R}(\Phi_2,\Phi_{+1})}{\mathrm{A}^{(0)}_{2\mapsto 3}(\Phi_{+1})\mathrm{B}(\Phi_2)}\,, \tag{9}$$

the following matching formula is NLO accurate up to terms appearing at order α_s^2

$$\langle O\rangle_{\text{NLO+PS}} = \int \mathrm{d}\Phi_2\, \mathrm{B}(\Phi_2)\, k_{\text{NLO}}(\Phi_2)\Big[\Delta(t_0,t_c)O(\Phi_2) + \int_{t_c}^{t_0} \mathrm{d}\Phi_{+1}\, w^{\text{LO}}_{2\mapsto 3}\mathrm{A}_{2\mapsto 3}(\Phi_{+1})\Delta(t_0,t)O(\Phi_2,\Phi_{+1})\Big]. \tag{10}$$

This can be seen by expanding the result to order α_s.

3. NNLO Matching Strategy

We now turn to the main result of this work, the definition of a fully-differential NNLO matching strategy for processes with two coloured final-state particles. It is applicable to both decays of colour singlets as well as scattering processes as long as all initial-state particles are colourless, for instance as in $e^+e^- \to jj$.

In the antenna formalism, the expected value for an infrared-safe observable of a process with two coloured final-state particles is given at NNLO by

$$\begin{aligned}\langle O\rangle_{\text{NNLO}} = \int \mathrm{d}\Phi_2 \Big[&\mathrm{B}(\Phi_2) + \mathrm{V}(\Phi_2) + \mathrm{I}^{\text{NLO}}_{\text{S}}(\Phi_2)\\ &+ \mathrm{VV}(\Phi_2) + \mathrm{I}_{\text{T}}(\Phi_2) + \mathrm{I}_{\text{S}}(\Phi_2)\Big] O(\Phi_2)\\ + \int \mathrm{d}\Phi_3 \Big[&\mathrm{R}(\Phi_2,\Phi_{+1})O(\Phi_2,\Phi_{+1})\\ &- \mathrm{S}^{\text{NLO}}(\Phi_2,\Phi_{+1})O(\Phi_2)\\ &+ \mathrm{RV}(\Phi_2,\Phi_{+1})O(\Phi_2,\Phi_{+1})\\ &+ \mathrm{T}(\Phi_2,\Phi_{+1},O)\Big]\\ + \int \mathrm{d}\Phi_4 \Big[&\mathrm{RR}(\Phi_2,\Phi_{+2})O(\Phi_2,\Phi_{+2})\\ &- \mathrm{S}(\Phi_2,\Phi_{+2},O)\Big],\end{aligned} \tag{11}$$

where RR is the differential double-real radiation cross section and RV and VV denote the differential virtual (one-loop) and double-virtual (two-loop) corrections to the real-radiation cross section R and the Born cross section B, respectively. In this context, the term S^{NLO} denotes the differential NLO real antenna subtraction term, S denotes the differential NNLO double-real antenna subtraction term [44–47],

$$\begin{aligned}\mathrm{S}(\Phi_2,\Phi_{+2},O) = \mathrm{S}^a(\Phi_2,\Phi_{+2})O(\Phi_2) + \mathrm{S}^b(\Phi_3,\Phi_{+1})O(\Phi_3)\\ - \mathrm{S}^c(\Phi_2,\Phi_{+1},\Phi'_{+1})O(\Phi_2)\end{aligned} \tag{12}$$

and T denotes the differential NNLO real-virtual antenna subtraction term [44–47],

$$\begin{aligned}\mathrm{T}(\Phi_2,\Phi_{+1},O) = \mathrm{T}^a(\Phi_2,\Phi_{+1})O(\Phi_2) + \mathrm{T}^b(\Phi_3)O(\Phi_3)\\ - \mathrm{T}^c(\Phi_2,\Phi_{+1})O(\Phi_2)\end{aligned} \tag{13}$$

Their integrated counterparts are given by $\mathrm{I}^{\text{NLO}}_{\text{S}}$, I_{T}, and I_{S}. In this context, terms labelled with superscript a constitute the double-real/real-virtual subtraction terms with compensating terms labelled with a superscript c that remove spurious single-unresolved singularities. The single-unresolved singularities are captured by the NLO subtraction terms of the +1-jet calculation, labelled with superscript b,

$$\mathrm{S}^b(\Phi_3,\Phi'_{+1}) \equiv \mathrm{S}^{\text{NLO}}(\Phi_3,\Phi'_{+1})\,, \quad \mathrm{T}^b(\Phi_3) \equiv \mathrm{I}^{\text{NLO}}_{\text{S}}(\Phi_3)\,. \tag{14}$$

The terms labeled with superscript b and c cancel independently in eq. (11). They are constructed such as to make the integrals individually infrared finite and thus amenable to evaluation with Monte-Carlo methods.

As for the NLO case, we define a Born-local weight,

$$
\begin{aligned}
k_{\mathrm{NNLO}}(\Phi_2) := 1 &+ \frac{\mathrm{V}(\Phi_2)}{\mathrm{B}(\Phi_2)} + \frac{\mathrm{I}_{\mathrm{S}}^{\mathrm{NLO}}(\Phi_2)}{\mathrm{B}(\Phi_2)} \\
&+ \frac{\mathrm{VV}(\Phi_2)}{\mathrm{B}(\Phi_2)} + \frac{\mathrm{I}_{\mathrm{T}}(\Phi_2)}{\mathrm{B}(\Phi_2)} + \frac{\mathrm{I}_{\mathrm{S}}(\Phi_2)}{\mathrm{B}(\Phi_2)} \\
&+ \int \mathrm{d}\Phi_{+1} \left[\frac{\mathrm{R}(\Phi_2, \Phi_{+1})}{\mathrm{B}(\Phi_2)} - \frac{\mathrm{S}^{\mathrm{NLO}}(\Phi_2, \Phi_{+1})}{\mathrm{B}(\Phi_2)} \right. \\
&\qquad \left. + \frac{\mathrm{RV}(\Phi_2, \Phi_{+1})}{\mathrm{B}(\Phi_2)} + \frac{\mathrm{T}(\Phi_2, \Phi_{+1})}{\mathrm{B}(\Phi_2)} \right] \\
&+ \int \mathrm{d}\Phi_{+2} \left[\frac{\mathrm{RR}(\Phi_2, \Phi_{+2})}{\mathrm{B}(\Phi_2)} - \frac{\mathrm{S}(\Phi_2, \Phi_{+2})}{\mathrm{B}(\Phi_2)} \right],
\end{aligned}
\tag{15}
$$

which will be used to construct the NNLO matching formula, and which can be used to perform the fixed-order computation in complete analogy to eq. (4). Here, $\mathrm{d}\Phi_{+2}$ is the two-particle radiation phase space that enters the factorised $n + 2$-particle phase space

$$
\mathrm{d}\Phi_{n+2} = \mathrm{d}\Phi_n \times \mathrm{d}\Phi_{+2} \,. \tag{16}
$$

We shall further need to distinguish between an ordered and unordered component of the two-particle radiation phase space, according to the following partition of unity:

$$
\begin{aligned}
\mathrm{d}\Phi_{+2} &= \theta(t' - t)\, \mathrm{d}\Phi_{+2} + \theta(t - t')\, \mathrm{d}\Phi_{+2} \,, \\
&= \mathrm{d}\Phi_{+2}^{>} + \mathrm{d}\Phi_{+2}^{<} \,.
\end{aligned}
\tag{17}
$$

The ordered part $\mathrm{d}\Phi_{+2}^{<}$ corresponds to the region accessible to strongly-ordered shower paths $t_0 > t > t'$, whereas the unordered part $\mathrm{d}\Phi_{+2}^{>}$ is inaccessible to strongly-ordered showers because of the larger intermediate scale $t_0 > t' > t$. We will use VINCIA's sector criterion, cf. sec. 3.3 in [33], to distinguish between the two, cf. section 4.2.

In order to be able to match the NNLO calculation with the shower, the shower needs to incorporate virtual corrections to ordinary $2 \to 3$ branchings as well as new $2 \to 4$ branchings, accounting for the simultaneous emission of two particles. These new shower terms correspond to the real-virtual and double-real corrections in the NNLO calculation. In addition, we need to incorporate the corresponding parton-shower counterterms. We start by defining the two-particle NLO Sudakov as [4]

$$
\begin{aligned}
&\Delta_2^{\mathrm{NLO}}(t_0, t) \\
&\quad = \exp\left\{ - \int_t^{t_0} \mathrm{d}\Phi_{+1}\, \mathrm{A}_{2\mapsto 3}^{(0)}(\Phi_{+1}) w_{2\mapsto 3}^{\mathrm{NLO}}(\Phi_2, \Phi_{+1}) \right\} \\
&\quad \times \exp\left\{ - \int_t^{t_0} \mathrm{d}\Phi_{+2}^{>}\, \mathrm{A}_{2\mapsto 4}^{(0)}(\Phi_{+2}) w_{2\mapsto 4}^{\mathrm{LO}}(\Phi_2, \Phi_{+2}) \right\},
\end{aligned}
\tag{18}
$$

where we have introduced the $2 \mapsto 4$ LO matrix-element correction factor,

$$
w_{2\mapsto 4}^{\mathrm{LO}}(\Phi_2, \Phi_{+2}) = \frac{\mathrm{RR}(\Phi_2, \Phi_{+2})}{\mathrm{A}_{2\mapsto 4}^{(0)}(\Phi_{+2}) \mathrm{B}(\Phi_2)} \tag{19}
$$

and the $2 \mapsto 3$ NLO matrix-element correction factor $w_{2\mapsto 3}^{\mathrm{NLO}}(\Phi_{+1})$, which we write in terms of a second order correction to the LO $2 \mapsto 3$ MEC in eq. (9),

$$
\begin{aligned}
w_{2\mapsto 3}^{\mathrm{NLO}}(\Phi_2, \Phi_{+1}) &= w_{2\mapsto 3}^{\mathrm{LO}}(\Phi_2, \Phi_{+1}) \\
&\quad \times \left(1 + \tilde{w}_{2\mapsto 3}^{\mathrm{FO}}(\Phi_2, \Phi_{+1}) + \tilde{w}_{2\mapsto 3}^{\mathrm{PS}}(\Phi_2)\right).
\end{aligned}
\tag{20}
$$

The coefficients $\tilde{w}$ are given by matching the $O\left(\alpha_{\mathrm{s}}^2\right)$ terms in the expansion of the truncated shower approximation to the fixed-order result in eq. (11) [4, 48]. We find the fixed-order contribution

$$
\begin{aligned}
&\tilde{w}_{2\mapsto 3}^{\mathrm{FO}}(\Phi_2, \Phi_{+1}) = \\
&\quad \frac{\mathrm{RV}(\Phi_2, \Phi_{+1})}{\mathrm{R}(\Phi_2, \Phi_{+1})} + \int_0^{t} \mathrm{d}\Phi'_{+1}\, \frac{\mathrm{RR}(\Phi_2, \Phi_{+1}, \Phi'_{+1})}{\mathrm{R}(\Phi_2, \Phi_{+1})} \\
&\qquad - \left(\frac{\mathrm{V}(\Phi_2)}{\mathrm{B}(\Phi_2)} + \int_0^{t_0} \mathrm{d}\Phi'_{+1}\, \frac{\mathrm{R}(\Phi_2, \Phi'_{+1})}{\mathrm{B}(\Phi_2)} \right),
\end{aligned}
\tag{21}
$$

and the second-order parton-shower matching term

$$
\begin{aligned}
\tilde{w}_{2\mapsto 3}^{\mathrm{PS}}(\Phi_2) &= \frac{\alpha_{\mathrm{s}}}{2\pi} \ln \frac{\kappa^2 \mu_{\mathrm{S}}^2}{\mu_{\mathrm{R}}^2} \\
&\quad + \int_t^{t_0} \mathrm{d}\Phi'_{+1}\, \mathrm{A}_{2\mapsto 3}^{(0)}(\Phi'_{+1}) w_{2\mapsto 3}^{\mathrm{LO}}(\Phi_2, \Phi'_{+1}) \,.
\end{aligned}
\tag{22}
$$

The factor κ is a constant and μ_{S}^2 is the parton-shower renormalisation scale. The two are conventionally chosen such that the logarithmic structure of eq. (21) is reproduced, which leads to $\mu_{\mathrm{S}} = p_\perp$ and $\kappa^2 = \exp\{K/\beta_0\}$, with K the two-loop cusp anomalous dimension [49–52]. This is known as the CMW scheme [53].

Note that in eq. (18), the integral over $\mathrm{A}_{2\mapsto 4}^{(0)}$ is defined over the range $[t, t_0]$, since the "ordered" contribution $t' < t$ has been reabsorbed into $\tilde{w}_{2\mapsto 3}^{\mathrm{FO}}(\Phi_{+1})$.

It should be emphasised that we do not require the NLO three-jet calculation to be provided externally but include the correction directly in the shower evolution. This means that, different to the situation in traditional merging approaches, this correction is exponentiated into a Sudakov factor. Up to the first emission, this agrees with the treatment in [5, 7] and implicitly includes the contribution from higher-order matching terms and collinear mass factorization counterterms that are needed to recover the NLO DGLAP splitting functions.

In addition, we need the 3-particle Sudakov, which we describe at LO,

$$\Delta_3^{\mathrm{LO}}(t,t') = \exp\left\{-\int_{t'}^{t} \mathrm{d}\Phi'_{+1}\, \mathrm{A}^{(0)}_{3\mapsto 4}(\Phi'_{+1}) w^{\mathrm{LO}}_{3\mapsto 4}(\Phi_3,\Phi'_{+1})\right\}. \tag{23}$$

with the $3 \mapsto 4$ LO matrix-element correction factor,

$$w^{\mathrm{LO}}_{3\mapsto 4}(\Phi_3,\Phi'_{+1}) = \frac{\mathrm{RR}(\Phi_2,\Phi_{+2})}{\mathrm{A}^{(0)}_{3\mapsto 4}(\Phi'_{+1})\mathrm{R}(\Phi_2,\Phi_{+1})}\,. \tag{24}$$

Up to the second emission, the shower operator is thus given by

$$\begin{aligned}\mathcal{S}_2(t_0,O) &= \Delta_2^{\mathrm{NLO}}(t_0,t_{\mathrm{c}})O(\Phi_2) \\ &+ \int_{t_{\mathrm{c}}}^{t_0} \mathrm{d}\Phi_{+1}\, \mathrm{A}^{(0)}_{2\mapsto 3}(\Phi_{+1}) w^{\mathrm{NLO}}_{2\mapsto 3} \Delta_2^{\mathrm{NLO}}(t_0,t) \\ &\quad\times\Big(\Delta_3^{\mathrm{LO}}(t,t_{\mathrm{c}})O(\Phi_2,\Phi_{+1}) \\ &\qquad+\int_{t_{\mathrm{c}}}^{t} \mathrm{d}\Phi'_{+1}\, \mathrm{A}^{(0)}_{3\mapsto 4}(\Phi'_{+1}) w^{\mathrm{LO}}_{3\mapsto 4}(\Phi_3,\Phi'_{+1}) O(\Phi_3,\Phi'_{+1})\Big) \\ &+ \int_{t_{\mathrm{c}}}^{t_0} \mathrm{d}\Phi^{>}_{+2}\, \mathrm{A}^{(0)}_{2\mapsto 4}(\Phi_{+2}) w^{\mathrm{LO}}_{2\mapsto 4}(\Phi_2,\Phi_{+2}) O(\Phi_2,\Phi_{+2})\end{aligned} \tag{25}$$

and our final NNLO+PS matching formula takes the simple form:

$$\langle O\rangle_{\mathrm{NNLO+PS}} = \int \mathrm{d}\Phi_2\, \mathrm{B}(\Phi_2) k_{\mathrm{NNLO}}(\Phi_2) \mathcal{S}_2(t_0,O)\,. \tag{26}$$

When expanding the truncated shower operator $\mathcal{S}_2$ in eq. (26) up to order α_{s}^2, NNLO accuracy is recovered for the observable $O(\Phi_2)$, while $O(\Phi_3)$ and $O(\Phi_4)$ achieve NLO and LO accuracy, respectively. This is true, because the combination of the iterated $2 \mapsto 3 \mapsto 4$ and the direct $2 \mapsto 4$ contributions to eq. (25) yields the correct double-real correction RR in eq. (11) by means of the LO MEC factors in eqs. (9), (19) and (24). Moreover the NLO correction eq. (20) recovers the correct real and real-virtual corrections R and RV in eq. (11) by means of eq. (9) and eq. (21).

4. Numerical Implementation

In this section, we want to present all necessary components of an implementation of our NNLO matching strategy. These are:

- a framework to calculate the Born-local NNLO K-factors in Eq. (15)
- a shower filling the strongly-ordered [54] and unordered [4] regions of the single- and double-emission phase space
- tree-level MECs in strongly-ordered [55] and unordered [56] shower paths
- NLO MECs in the first emission [48]

With the exception of the first point, (process-dependent) implementations of these components existed in previous VINCIA versions (not necessarily simultaneously), and have been described in detail in the various references. We have (re-)implemented all components in a semi-automated [3] fashion in the VINCIA antenna shower in PYTHIA 8.3. We access loop matrix elements via a novel MCFM [57–60] interface presented in [61] and tree-level matrix elements via a new run-time interface [62] to the COMIX matrix element generator [63] in SHERPA [64, 65].

Our NNLO matching algorithm can be summarised in the following steps:

1. Generate a phase space point according to the Born cross section $\mathrm{B}(\Phi_2)$.
2. Calculate the Born-local NNLO factor $k_{\mathrm{NNLO}}(\Phi_2)$ and reweight the phase space point by the result.
3. Let the phase-space maximum given by the invariant mass of the two Born partons define the starting scale for the shower, $t_{\mathrm{now}} = t_0(\Phi_2)$.
4. Starting from the current shower scale, t_{now}, let the $2 \mapsto 3$ and $2 \mapsto 4$ showers compete for the highest branching scale.
5. Update the current shower scale to be that of the winning branching, $t_{\mathrm{now}} = \max(t_{2\mapsto 3}, t_{2\mapsto 4})$.

6a. If the winning branching is a $2 \mapsto 3$ branching, calculate the accept probability including the NLO MEC $w^{\mathrm{NLO}}_{2\mapsto 3}$.
   - If rejected, continue from step 4.
   - If accepted, continue with a LO shower from the resulting three-particle configuration, starting from t_{now} and including the LO MEC $w^{\mathrm{LO}}_{3\mapsto 4}$ when calculating accept probabilities for the $3 \mapsto 4$ step.

   When a $3 \mapsto 4$ branching is accepted (or the shower cutoff scale is reached), continue with step 7.

6b. If the winning branching is a $2 \mapsto 4$ branching, calculate the accept probability including the LO MEC $w^{\mathrm{LO}}_{2\mapsto 4}$.

[3] Semi-automated here refers to the fact that antenna subtraction terms are explicitly implemented for each class of processes.

- If rejected, continue from step 4.
- If accepted, continue with step 7.

7. Continue with a standard (possibly uncorrected) shower from the resulting four-particle configuration, starting from t_{now}.

It should be emphasised that the matrix-element correction factors make this algorithm independent of the splitting kernels (i.e. antenna functions in our case) up to the matched order and the shower merely acts as an efficient Sudakov-weighted phase-space generator. Hence, if the algorithm is stopped after step 6, an NNLO-matched result is obtained, which can be showered by any other parton shower, just as is the case for POWHEG NLO matching. Note, that there remains a dependence on the ordering variable, which has to be properly accounted for.

4.1. NNLO Kinematics

For both, the unordered shower contributions and the Born-local NNLO weight, new kinematic maps are needed to reflect their direct $2 \mapsto 4$, i.e. unordered or double-unresolved, nature. We utilise that the n-particle phase space measure may be factorised into the product of a $2 \mapsto 3$ antenna phase space and the $n-1$-particle phase space measure, as well as into the product of a $2 \mapsto 4$ antenna phase space and the $n-2$-particle phase space. This allows us to write the $2 \mapsto 4$ antenna phase space as the product of two $2 \mapsto 3$ antenna phase spaces,

$$\begin{aligned} &\mathrm{d}\Phi_{+2}(p_I + p_K; p_i, p_{j_1}, p_{j_2}, p_k) \\ &\qquad = \mathrm{d}\Phi_{+1}(p_I + p_K; \hat{p}_i, \hat{p}_j, p_k) \\ &\qquad\quad \times \mathrm{d}\Phi_{+1}(\hat{p}_i + \hat{p}_j; p_i, p_{j_1}, p_{j_2}), \end{aligned} \tag{27}$$

corresponding to the kinematic mapping

$$p_I + p_K = \hat{p}_i + \hat{p}_j + p_k = p_i + p_{j_1} + p_{j_2} + p_k, \tag{28}$$

effectively representing a tripole map [66]. In line with the phase space factorisation, the kinematic mapping is then constructed as an iteration of two on-shell $2 \mapsto 3$ antenna maps given in sec. 2.3 in [33].

We have tested the validity of our kinematic maps by comparing VINCIA's phase-space mappings (double-gluon emission and gluon-emission-plus-splitting) to a flat sampling via RAMBO.

4.2. Unordered Shower Contributions

An important part of our proposal is the inclusion of double-unresolved radiation in the shower evolution. To this end, we employ the sector-antenna framework [33] and amend it by direct $2 \mapsto 4$ branchings as described in [4]. In the sector-shower approach, each branching is restricted to the region in phase space where it minimises the resolution variable, defined for final-state clusterings by

$$Q^2_{\text{res},j} = \begin{cases} \frac{s_{ij}s_{jk}}{s_{IK}} & \text{if } j \text{ is a gluon} \\ s_{ij}\sqrt{\frac{s_{jk}}{s_{IK}}} & \text{if } (i,j) \text{ is a quark-antiquark pair} \end{cases} \tag{29}$$

This is achieved by a "sectorisation" of phase space according the partition of unity,

$$1 = \sum_j \Theta^{\text{sct}}_{j/IK} = \sum_j \theta\left(\min_i \left\{Q^2_{\text{res},i}\right\} - Q^2_{\text{res},j}\right), \tag{30}$$

which is implemented in the shower evolution as an explicit veto for each trial branching. Since only a single branching kernel contributes per colour-ordered phase space point, sector antenna functions have to incorporate the full singularity structure associated with the respective sector. At LO, this amounts to including both the full single-collinear and single-soft limits in the antenna function. The full set of VINCIA's LO sector antenna functions is collected in [33].

By construction, the default sector shower generates only strongly-ordered sequences[4], as the sector veto ensures that each emission is the softest (or most-collinear) in the post-branching configuration. The inclusion of direct $2 \mapsto 4$ branchings (which look unordered from an iterated $2 \mapsto 3$ point of view) in the sector shower is facilitated by extending the sector decomposition in eq. (30) by an ordering criterion,

$$\begin{aligned} 1 &= \sum_j \left[\Theta^{<}_{j/IK}\Theta^{\text{sct}}_{j/IK} + \Theta^{>}_{j/IK}\Theta^{\text{sct}}_{j/IK}\right] \\ &= \underbrace{\sum_j \theta\left(\hat{p}^2_{\perp,\hat{j}} - p^2_{\perp,j}\right)\Theta^{\text{sct}}_{j/IK}}_{2\mapsto 3 \text{ (strongly ordered)}} + \underbrace{\sum_j \theta\left(p^2_{\perp,j} - \hat{p}^2_{\perp,\hat{j}}\right)\Theta^{\text{sct}}_{j/IK}}_{2\mapsto 4 \text{ (unordered)}} \end{aligned} \tag{31}$$

where $p^2_\perp$ denotes VINCIA's transverse-momentum ordering variable and hatted variables denote the intermediate node in a sequence $IL \mapsto \hat{i}\hat{j}\hat{\ell} \mapsto ijk\ell$. Here, the scales $p^2_\perp$ and $\hat{p}^2_\perp$ are uniquely defined by the ordering variable of the sector-shower emission, i.e., that emission which minimises eq. (29). Direct $2 \mapsto 4$ emissions are thus restricted to the unordered region of the double-emission phase space, denoted as $\mathrm{d}\Phi^{>}_{+2}$ in eq. (18) and

[4] This is different to virtually any other strongly-ordered shower, where recoil effects introduce unordered sequences. Such phase space points are vetoed in a sector shower.

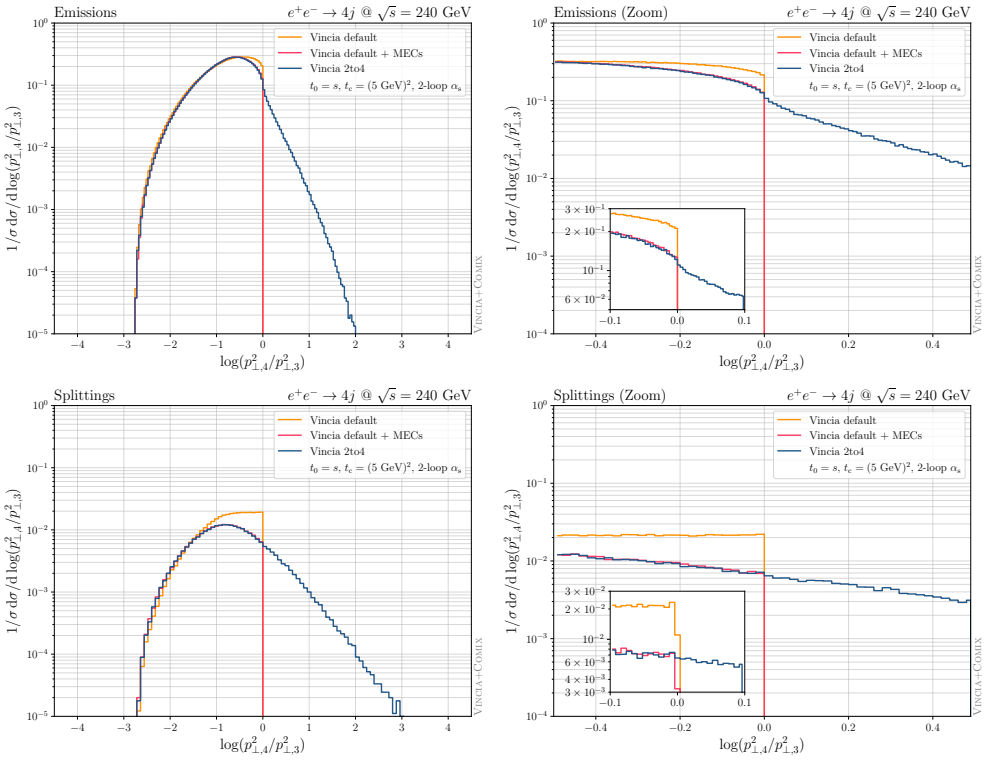


Figure 1: Ratio of the evolution variable of the four-parton and three-parton configuration $\log\left(p^2_{\perp,4}/p^2_{\perp,3}\right)$ in $e^+e^- \to 4j$. The region > 0 corresponds to unordered contributions not reached by strongly-ordered showers.

defined as

$$\mathrm{d}\Phi^{>}_{+2} = \sum_j \Theta^{>}_{j/IK}\Theta^{\mathrm{sct}}_{j/IK}\,\mathrm{d}\Phi^{j}_{+2}\,. \tag{32}$$

For $2 \to 4$ emissions off quark-antiquark and gluon-gluon antennae, we use the double-real antenna functions in [44, 45, 47]. We note that NLO quark-gluon antenna functions appear in the Standard Model at lowest order for three final-state particles and are hence not of interest for our test case of $e^+e^- \to jj$. We wish to point out, however, that the NLO quark-gluon antenna functions in [46, 47] contain spurious singularities which have to be removed before a shower implementation is possible.

As a validation, we show in fig. 1 the ratio of the four-jet to three-jet evolution variable for $e^+e^- \to 4j$ at $\sqrt{s} = 240$ GeV. To focus on the perturbative realm, the shower evolution is constrained to the region between $t_0 = s$ and $t_c = (5\ \mathrm{GeV})^2$. The region > 0 corresponds to the unordered part of phase space to which strongly-ordered showers cannot contribute. Due to the use of sector showers, there is a sharp cut-off at the boundary between the ordered and unordered region, as the sector criterion ensures that the last emission is always the softest and therefore, no recoil effects can spoil the strong ordering of the shower. As expected, the inclusion of direct $2 \to 4$ branchings gives access to the unordered parts of phase space, a crucial element of our matching method.

4.3. LO Matrix-Element Corrections

In order for the shower expansion to match the fixed-order calculation, we need (iterated) $2 \mapsto 3$ tree-level MECs and (direct) $2 \mapsto 4$ tree-level MECs. Both take a particularly simple form in the sector-antenna framework, as will be shown below.

At leading-colour, tree-level MECs to the ordered sector shower can be constructed as [55, 67]

$$w^{\mathrm{LO,LC}}_{2\mapsto 3,i}(\Phi_2, \Phi_{+1}) = \frac{\mathrm{R}^{\mathrm{LC}}_i(\Phi_2, \Phi_{+1})}{\sum_j \Theta^{\mathrm{sct}}_{j/IK} A^{\mathrm{sct}}_{j/IK}(p_i, p_j, p_k)\mathrm{B}(\Phi_2)}\,,$$

$$w^{\mathrm{LO,LC}}_{3\mapsto 4,i}(\Phi_3, \Phi_{+1}) = \frac{\mathrm{RR}^{\mathrm{LC}}_i(\Phi_3, \Phi_{+1})}{\sum_j \Theta^{\mathrm{sct}}_{j/IK} A^{\mathrm{sct}}_{j/IK}(p_i, p_j, p_k)\mathrm{R}^{\mathrm{LC}}_i(\Phi_3)}\,,$$

where

$$\mathrm{B}(\Phi_2) = \left|\mathcal{M}_2^{(0)}(p_1, p_2)\right|^2 ,$$
$$\mathrm{R}_i^{\mathrm{LC}}(\Phi_3) = \left|\mathcal{M}_3^{(0)}(\sigma_i\{p_1, p_2, p_3\})\right|^2 ,$$
$$\mathrm{RR}_i^{\mathrm{LC}}(\Phi_4) = \left|\mathcal{M}_4^{(0)}(\sigma_i\{p_1, p_2, p_3, p_4\})\right|^2 ,$$

denote squared leading-colour colour-ordered amplitudes with the index i denoting the respective permutation σ_i (the number of permutations depends on the process). The sector veto $\Theta^{\mathrm{sct}}_{j/IK}$ ensures that only the most singular term contributes in the denominators, rendering the fraction exceptionally simple.

Direct $2 \mapsto 4$ branchings can be corrected in an analogous way, replacing the sum over $2 \mapsto 3$ antenna functions with a sum of $2 \mapsto 4$ ones,

$$w^{\mathrm{LO,LC}}_{2\mapsto 4,i}(\Phi_2, \Phi_{+2}) = \frac{\mathrm{RR}_i^{\mathrm{LC}}(\Phi_2, \Phi_{+2})}{\sum_{\{j,k\}} \Theta^{\mathrm{sct}}_{jk/IL} A^{\mathrm{sct}}_{jk/IL}(p_i, p_j, p_k, p_\ell)\mathrm{B}(\Phi_2)} ,$$

The full-colour matrix element can be recovered on average by multiplication with a full-colour to leading-colour-summed matrix-element weight,

$$w^{\mathrm{LO}}_{2\mapsto 3,i} = w^{\mathrm{LO,LC}}_{2\mapsto 3,i} \times \frac{\mathrm{R}(\Phi_2, \Phi_{+1})}{\sum_j \mathrm{R}_j^{\mathrm{LC}}(\Phi_2, \Phi_{+1})} , \tag{33}$$

$$w^{\mathrm{LO}}_{3\mapsto 4,i} = w^{\mathrm{LO,LC}}_{3\mapsto 4,i} \times \frac{\mathrm{RR}(\Phi_3, \Phi_{+1})}{\sum_j \mathrm{RR}_j^{\mathrm{LC}}(\Phi_3, \Phi_{+1})} , \tag{34}$$

$$w^{\mathrm{LO}}_{2\mapsto 4,i} = w^{\mathrm{LO,LC}}_{2\mapsto 4,i} \times \frac{\mathrm{RR}(\Phi_2, \Phi_{+2})}{\sum_j \mathrm{RR}_j^{\mathrm{LC}}(\Phi_2, \Phi_{+2})} . \tag{35}$$

For gluon splittings, multiple histories contribute even in the sector shower, because all permutations of quark lines have to be taken into account. To ensure that the MEC factors remain finite for final states with multiple quark pairs, an additional quark-projection factor has to be included. Since we only deal with a maximum of two quark pairs, it is given by

$$\rho_j = \frac{A^{\mathrm{sct}}_{j_q/g_I X_K}(\bar{q}_i, q_j, X_k)}{\sum_j A^{\mathrm{sct}}_{j_q/g_I X_K}(\bar{q}_i, q_j, X_k)} \tag{36}$$

for $2 \to 3$ branchings and

$$\rho_j = \frac{A^{\mathrm{sct}}_{j_q k_{\bar{q}}/X_I Y_L}(X_i, q_j, \bar{q}_k, Y_\ell)}{\sum_j A^{\mathrm{sct}}_{j_q k_{\bar{q}}/X_I Y_L}(X_i, q_j, \bar{q}_k, Y_\ell)} \tag{37}$$

for $2 \mapsto 4$ branchings.

4.4. NLO Matrix-Element Corrections

Making the antenna subtraction terms explicit, the fixed-order correction to the NLO matrix-element correction eq. (20) reads

$$\begin{aligned}\tilde{w}^{\mathrm{FO}}_{2\mapsto 3}(\Phi_2, \Phi_{+1}) = {} & \frac{\mathrm{RV}(\Phi_2, \Phi_{+1})}{\mathrm{R}(\Phi_2, \Phi_{+1})} + \frac{\mathrm{I}^{\mathrm{NLO}}(\Phi_2, \Phi_{+1})}{\mathrm{R}(\Phi_2, \Phi_{+1})} \\ & + \int_0^t \mathrm{d}\Phi'_{+1} \left[\frac{\mathrm{RR}(\Phi_2, \Phi_{+1}, \Phi'_{+1})}{\mathrm{R}(\Phi_2, \Phi_{+1})} - \frac{\mathrm{S}^{\mathrm{NLO}}(\Phi_2, \Phi_{+1}, \Phi'_{+1})}{\mathrm{R}(\Phi_2, \Phi_{+1})}\right] \\ & - \left(\frac{\mathrm{V}(\Phi_2)}{\mathrm{B}(\Phi_2)} + \frac{\mathrm{I}^{\mathrm{NLO}}(\Phi_2)}{\mathrm{B}(\Phi_2)} \right. \\ & \left. + \int_0^{t_0} \mathrm{d}\Phi'_{+1} \left[\frac{\mathrm{R}(\Phi_2, \Phi'_{+1})}{\mathrm{B}(\Phi_2)} - \frac{\mathrm{S}^{\mathrm{NLO}}(\Phi_2, \Phi'_{+1})}{\mathrm{B}(\Phi_2)}\right]\right),\end{aligned} \tag{38}$$

with the differential NLO antenna subtraction terms $\mathrm{S}^{\mathrm{NLO}}(\Phi_2, \Phi'_{+1})$, $\mathrm{S}^{\mathrm{NLO}}(\Phi_2, \Phi_{+1}, \Phi'_{+1})$ and their integrated counterparts $\mathrm{I}_{\mathrm{S}}^{\mathrm{NLO}}(\Phi_2)$, $\mathrm{I}_{\mathrm{S}}^{\mathrm{NLO}}(\Phi_2, \Phi_{+1})$ cf. eqs. (11) and (14). Based on the argument of the last subsection, we construct the full-colour NLO matrix-element correction as

$$\begin{aligned}w^{\mathrm{NLO}}_{2\mapsto 3,i}(\Phi_2, \Phi_{+1}) = {} & w^{\mathrm{LO,LC}}_{2\mapsto 3,i}(\Phi_2, \Phi_{+1}) \frac{\mathrm{R}(\Phi_2, \Phi_{+1})}{\sum_j \mathrm{R}_j^{\mathrm{LC}}(\Phi_2, \Phi_{+1})} \\ & \times (1 + \tilde{w}^{\mathrm{FO}}_{2\mapsto 3}(\Phi_2, \Phi_{+1}) + \tilde{w}^{\mathrm{PS}}_{2\mapsto 3}(\Phi_2)) .\end{aligned} \tag{39}$$

The integration over the radiation phase spaces denoted Φ'_{+1} in eq. (38) is done numerically, utilising antenna kinematics to map 3-parton configurations to 4-parton configurations (similarly for 2-parton configurations). This phase-space generation approach will be described in detail in the next subsection in the context of the NNLO Born weight. Note that the radiation phase space Φ_{+1} in eq. (38) is generated by the shower.

4.5. NNLO Born Weight

The Born-local NNLO weight can be calculated numerically using a "forward-branching" phase-space generation approach [36, 37, 68, 69, 71], which has previously been applied to unweighted NLO event generation, using Catani-Seymour dipole subtraction [73]. The application to NNLO corrections to $e^+e^- \to 2j$ using antenna subtraction has been outlined in [74].

Given a Born phase space point, the real-radiation phase space is generated by uniformly sampling the shower variables (t, ζ, ϕ) for each antenna, which represent integration channels in this context. As for the shower evolution, every phase space point is restricted to the sector in which the emission(s) correspond to the most-singular clusterings. The momenta

of the Born+1j point are constructed according to the same kinematic map as the shower uses, summarised in sec. 2.3 in [54]. Since antenna functions are azimuthally averaged, they do not cancel spin-correlations in collinear gluon branchings locally. To obtain a pointwise pole cancellation, the subtracted real correction R − S can be evaluated on two correlated phase space points,

$$\left\{(t, \zeta, \phi), (t, \zeta, \phi + \pi/2)\right\}$$

which cancels the collinear spin correlation exactly, as it is proportional to $\cos(2\phi)$. To obtain double-real radiation phase space points for the subtracted double-real correction RR − S, this procedure can be iterated, yielding four angular-correlated phase space points which cancel spin correlations in double single-collinear and triple-collinear limits. Due to the bijective nature of the sector-antenna framework, each 3- or 4-particle phase-space point obtained in this way can be mapped back uniquely to its 2-particle origin, making the NNLO weight exactly Born-local. For $e^+e^- \to 2j$ this procedure is identical to the one in [74].

We have implemented the NNLO antenna subtraction terms for processes with two massless final-state jets, cf. e.g. [44], in VINCIA in a semi-automated fashion. As a validation, we illustrate the convergence of the double-real radiation subtraction term eq. (12) in the triple-collinear and double-soft limits for the process $e^+e^- \to qgg\bar{q}$ in fig. 2. Phase space points are sampled according to the kinematic map in section 4.1 and we do not make use of the azimuthal averaging alluded to above.

It should be noted that a numerical calculation of the Born-local NNLO weight is not necessary for colour-singlet decays, as the inclusive K-factors are well known from analytical calculations, cf. e.g. [44, 75] for $Z \to q\bar{q}$ (with massless quarks), [76–79] for $H \to b\bar{b}$ (with massless bs), and [45, 80] for $H \to gg$ (in the Higgs effective theory).

5. Conclusions and Outlook

We have presented a technique to match final-state parton showers fully-differentially to next-to-next-to-leading order calculations in processes with two final-state jets. To our knowledge, this is the first method of its kind.

We have outlined a full-fledged numerical implementation in the VINCIA antenna shower in the PYTHIA 8.3 event generator. Phenomenological studies employing our strategy will be presented in separate works.

We want to close by noting that, while we here focused on the simplest case of two massless final-state jets, the use of the NNLO antenna subtraction formalism facilitates its adaption to more complicated processes such as $e^+e^- \to t\bar{t}$ or $e^+e^- \to 3j$. Considering the latter, spurious singularities in the quark-gluon NNLO antenna subtraction terms need to be removed before exponentiation in the shower. For future work, an extension of our method to processes with coloured initial states can be envisioned, given the applicability of the NNLO antenna subtraction to hadronic collisions.

Acknowledgements

We thank Aude Gehrmann-de Ridder and Thomas Gehrmann for providing us with FORM files of their antenna functions. We thank Philip Ilten for the development of a general matrix-element generator interface for PYTHIA 8.3, which allowed us to interface COMIX in this work. CTP is supported by the Monash Graduate Scholarship, the Monash International Postgraduate Research Scholarship, and the J.L. William Scholarship. HTL is supported by the U.S. Department of Energy under Contract No. DE-AC02-06CH11357 and the National Science Foundation under Grant No. NSF-1740142. This research was supported by Fermi National Accelerator Laboratory (Fermilab), a U.S. Department of Energy, Office of Science, HEP User Facility. Fermilab is managed by Fermi Research Alliance, LLC (FRA), acting under Contract No. DE-AC02-07CH11359. This work was further partly funded by the Australian Research Council via Discovery Project DP170100708 — "Emergent Phenomena in Quantum Chromodynamics". This work was also supported in part by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 722104 – MCnetITN3.

References

[1] S. Frixione, B. R. Webber, Matching NLO QCD computations and parton shower simulations, JHEP 06 (2002) 029. arXiv:hep-ph/0204244.

[2] P. Nason, A New method for combining NLO QCD with shower Monte Carlo algorithms, JHEP 11 (2004) 040. arXiv:hep-ph/0409146.

[3] S. Höche, F. Krauss, M. Schonherr, F. Siegert, A critical appraisal of NLO+PS matching methods, JHEP 09 (2012) 049. arXiv:1111.1220, doi:10.1007/JHEP09(2012)049.

[4] H. T. Li, P. Skands, A framework for second-order parton showers, Phys. Lett. B 771 (2017) 59–66. arXiv:1611.00013.

[5] S. Höche, S. Prestel, Triple collinear emissions in parton showers, Phys. Rev. D96 (7) (2017) 074017. arXiv:1705.00742.

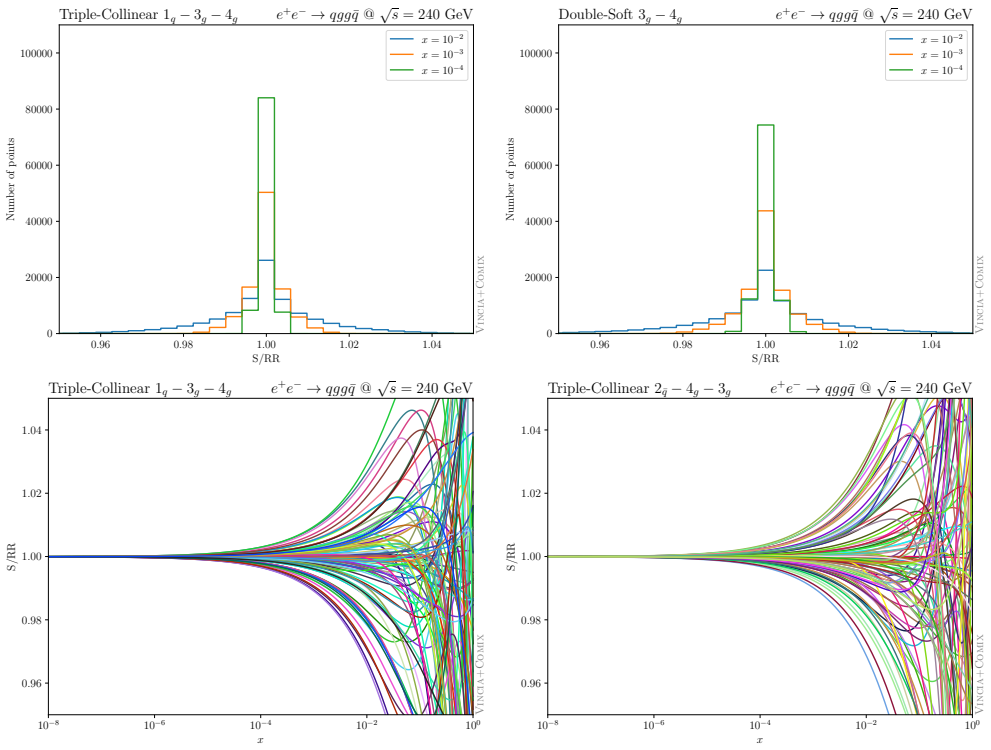


Figure 2: Test of the convergence of the double-real subtraction term $\mathrm{S}(\Phi_2, \Phi_{+2}, O)$ in eq. (12) in $e^+e^- \to qgg\bar{q}$. *Top row*: progression of weight distributions from $x = 10^{-2}$ to $x = 10^{-4}$ in the triple-collinear limit ($s_{134}/s_{1234} < x$) and double-soft limit ($s_{134}s_{234}/s^2_{1234} < x$). *Bottom row*: trajectories $x \cdot s_{134}$, $x \cdot s_{234}$, $x \to 0$ approaching the two triple collinear limits. Phase space points are not azimuthally averaged.

[6] S. Höche, F. Krauss, S. Prestel, Implementing NLO DGLAP evolution in Parton Showers, JHEP 10 (2017) 093. arXiv:1705.00982.

[7] F. Dulat, S. Höche, S. Prestel, Leading-Color Fully Differential Two-Loop Soft Corrections to QCD Dipole Showers, Phys. Rev. D98 (7) (2018) 074013. arXiv:1805.03757.

[8] S. Höche, Y. Li, S. Prestel, Drell-Yan lepton pair production at NNLO QCD with parton showers, Phys. Rev. D 91 (7) (2015) 074015. arXiv:1405.3607.

[9] L. Lönnblad, S. Prestel, Merging Multi-leg NLO Matrix Elements with Parton Showers, JHEP 03 (2013) 166. arXiv:1211.7278.

[10] P. F. Monni, P. Nason, E. Re, M. Wiesemann, G. Zanderighi, MiNNLO$_{PS}$: a new method to match NNLO QCD to parton showers, JHEP 05 (2020) 143. arXiv:1908.06987.

[11] E. Re, MiNNLO$_{\mathrm{PS}}$: a new method to match NNLO QCD with parton showers, in: 55th Rencontres de Moriond on QCD and High Energy Interactions, 2021. arXiv:2107.06753.

[12] K. Hamilton, P. Nason, E. Re, G. Zanderighi, NNLOPS simulation of Higgs boson production, JHEP 10 (2013) 222. arXiv:1309.0017.

[13] A. Karlberg, E. Re, G. Zanderighi, NNLOPS accurate Drell-Yan production, JHEP 09 (2014) 134. arXiv:1407.2940.

[14] W. Astill, W. Bizon, E. Re, G. Zanderighi, NNLOPS accurate associated HW production, JHEP 06 (2016) 154. arXiv:1603.01620.

[15] W. Astill, W. Bizoń, E. Re, G. Zanderighi, NNLOPS accurate associated HZ production with $H \to b\bar{b}$ decay at NLO, JHEP 11 (2018) 157. arXiv:1804.08141.

[16] K. Hamilton, P. Nason, C. Oleari, G. Zanderighi, Merging H/W/Z + 0 and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching, JHEP 05 (2013) 082. arXiv:1212.4504.

[17] R. Frederix, K. Hamilton, Extending the MINLO method, JHEP 05 (2016) 042. arXiv:1512.02663.

[18] Y. Hu, C. Sun, X.-M. Shen, J. Gao, Hadronic decays of Higgs boson at NNLO matched with parton shower (2021). arXiv:2101.08916.

[19] K. Hamilton, P. Nason, G. Zanderighi, MINLO: Multi-Scale Improved NLO, JHEP 10 (2012) 155. arXiv:1206.3572.

[20] S. Prestel, Matching N3LO QCD calculations to parton showers (2021). arXiv:2106.03206.

[21] S. Alioli, C. W. Bauer, C. Berggren, F. J. Tackmann, J. R. Walsh, S. Zuberi, Matching Fully Differential NNLO Calculations and Parton Showers, JHEP 06 (2014) 089. arXiv:1311.0286.

[22] S. Alioli, A. Broggio, A. Gavardi, S. Kallweit, M. A. Lim, R. Nagar, D. Napoletano, C. W. Bauer, L. Rottoli, Matching NNLO to parton shower using N^{3}LL colour-singlet transverse momentum resummation in GENEVA (2021). arXiv:2102.08390.

[23] S. Höche, Y. Li, S. Prestel, Higgs-boson production through gluon fusion at NNLO QCD with parton showers, Phys. Rev. D 90 (5) (2014) 054011. arXiv:1407.3773.

[24] S. Höche, S. Kuttimalai, Y. Li, Hadronic Final States in DIS at NNLO QCD with Parton Showers, Phys. Rev. D 98 (11) (2018) 114013. arXiv:1809.04192.

[25] S. Alioli, C. W. Bauer, C. Berggren, F. J. Tackmann, J. R. Walsh, Drell-Yan production at NNLL'+NNLO matched to par-

ton showers, Phys. Rev. D 92 (9) (2015) 094020. arXiv:1508.01475.

[26] S. Alioli, C. W. Bauer, S. Guns, F. J. Tackmann, Underlying event sensitive observables in Drell-Yan production using GENEVA, Eur. Phys. J. C 76 (11) (2016) 614. arXiv:1605.07192.

[27] T. Cridge, M. A. Lim, R. Nagar, $W\gamma$ production at NNLO+PS accuracy in GENEVA (2021). arXiv:2105.13214.

[28] P. F. Monni, E. Re, M. Wiesemann, MiNNLO$_{\rm PS}$: optimizing $2 \to 1$ hadronic processes, Eur. Phys. J. C 80 (11) (2020) 1075. arXiv:2006.04133.

[29] D. Lombardi, M. Wiesemann, G. Zanderighi, Advancing MiNNLO$_{\rm PS}$ to diboson processes: $Z\gamma$ production at NNLO+PS (2020). arXiv:2010.10478.

[30] D. Lombardi, M. Wiesemann, G. Zanderighi, W^+W^- production at NNLO+PS with MiNNLO$_{\rm PS}$ (2021). arXiv:2103.12077.

[31] J. Mazzitelli, P. F. Monni, P. Nason, E. Re, M. Wiesemann, G. Zanderighi, Next-to-next-to-leading order event generation for top-quark pair productionarXiv:2012.14267.

[32] L. Buonocore, G. Koole, D. Lombardi, L. Rottoli, M. Wiesemann, G. Zanderighi, ZZ production at nNNLO+PS with MiNNLO$_{\rm PS}$arXiv:2108.05337.

[33] H. Brooks, C. T. Preuss, P. Skands, Sector Showers for Hadron Collisions, JHEP 07 (2020) 032. arXiv:2003.00702.

[34] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, P. Skands, An Introduction to PYTHIA 8.2, Comput. Phys. Commun. 191 (2015) 159–177. arXiv:1410.3012.

[35] E. Norrbin, T. Sjöstrand, QCD radiation off heavy particles, Nucl. Phys. B 603 (2001) 297–342. arXiv:hep-ph/0010012.

[36] S. Frixione, P. Nason, C. Oleari, Matching NLO QCD computations with Parton Shower simulations: the POWHEG method, JHEP 11 (2007) 070. arXiv:0709.2092.

[37] S. Alioli, P. Nason, C. Oleari, E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX, JHEP 06 (2010) 043. arXiv:1002.2581.

[38] S. Frixione, Z. Kunszt, A. Signer, Three jet cross-sections to next-to-leading order, Nucl. Phys. B 467 (1996) 399–442. arXiv:hep-ph/9512328, doi:10.1016/0550-3213(96)00110-1.

[39] S. Catani, M. H. Seymour, A General algorithm for calculating jet cross-sections in NLO QCD, Nucl. Phys. B485 (1997) 291–419, [Erratum: Nucl. Phys.B510,503(1998)]. arXiv:hep-ph/9605323.

[40] D. A. Kosower, Antenna factorization of gauge theory amplitudes, Phys. Rev. D57 (1998) 5410–5416. arXiv:hep-ph/9710213.

[41] D. A. Kosower, Antenna factorization in strongly ordered limits, Phys. Rev. D71 (2005) 045016. arXiv:hep-ph/0311272.

[42] A. J. Larkoski, M. E. Peskin, Spin-Dependent Antenna Splitting Functions, Phys. Rev. D81 (2010) 054010. arXiv:0908.2450.

[43] J. J. Lopez-Villarejo, P. Z. Skands, Efficient Matrix-Element Matching with Sector Showers, JHEP 11 (2011) 150. arXiv:1109.3608, doi:10.1007/JHEP11(2011)150.

[44] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, Infrared structure of e+ e- —> 2 jets at NNLO, Nucl. Phys. B 691 (2004) 195–222. arXiv:hep-ph/0403057.

[45] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, Gluon-gluon antenna functions from Higgs boson decay, Phys. Lett. B 612 (2005) 49–60. arXiv:hep-ph/0502110.

[46] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, Quark-gluon antenna functions from neutralino decay, Phys. Lett. B 612 (2005) 36–48. arXiv:hep-ph/0501291.

[47] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, Antenna subtraction at NNLO, JHEP 09 (2005) 056. arXiv:hep-ph/0505111.

[48] L. Hartgring, E. Laenen, P. Skands, Antenna Showers with One-Loop Matrix Elements, JHEP 10 (2013) 127. arXiv:1303.4974.

[49] J. Kodaira, L. Trentadue, Summing Soft Emission in QCD, Phys. Lett. B 112 (1982) 66. doi:10.1016/0370-2693(82)90907-8.

[50] C. T. H. Davies, W. J. Stirling, Nonleading Corrections to the Drell-Yan Cross-Section at Small Transverse Momentum, Nucl. Phys. B 244 (1984) 337–348. doi:10.1016/0550-3213(84)90316-X.

[51] C. T. H. Davies, B. R. Webber, W. J. Stirling, Drell-Yan Cross-Sections at Small Transverse Momentum, Nucl. Phys. B 256 (1985) 413. doi:10.1016/0550-3213(85)90402-X.

[52] S. Catani, E. D'Emilio, L. Trentadue, The Gluon Formfactor to Higher Orders: Gluon Gluon Annihilation at Small Q^-transverse, Phys. Lett. B 211 (1988) 335–342. doi:10.1016/0370-2693(88)90912-4.

[53] S. Catani, B. R. Webber, G. Marchesini, QCD coherent branching and semiinclusive processes at large x, Nucl. Phys. B 349 (1991) 635–654. doi:10.1016/0550-3213(91)90390-J.

[54] H. Brooks, C. T. Preuss, Efficient Multi-Jet Merging at High Multiplicitiesarxiv:2008.09468.

[55] N. Fischer, S. Prestel, Combining states without scale hierarchies with ordered parton showers, Eur. Phys. J. C77 (9) (2017) 601. arXiv:1706.06218.

[56] W. T. Giele, D. A. Kosower, P. Z. Skands, Higher-Order Corrections to Timelike Jets, Phys. Rev. D84 (2011) 054003. arXiv:1102.2126.

[57] J. M. Campbell, R. K. Ellis, An Update on vector boson pair production at hadron colliders, Phys. Rev. D 60 (1999) 113006. arXiv:hep-ph/9905386.

[58] J. M. Campbell, R. K. Ellis, C. Williams, Vector boson pair production at the LHC, JHEP 07 (2011) 018. arXiv:1105.0020.

[59] J. M. Campbell, R. K. Ellis, W. T. Giele, A Multi-Threaded Version of MCFM, Eur. Phys. J. C 75 (6) (2015) 246. arXiv:1503.06182.

[60] J. Campbell, T. Neumann, Precision Phenomenology with MCFM, JHEP 12 (2019) 034. arXiv:1909.09117.

[61] J. M. Campbell, S. Höche, C. T. Preuss, Accelerating LHC phenomenology with analytic one-loop amplitudes: A C++ interface to MCFM (2021). arXiv:2107.04472.

[62] S. Höche, C. T. Preuss, In preparation.

[63] T. Gleisberg, S. Höche, Comix, a new matrix element generator, JHEP 12 (2008) 039. arXiv:0808.3674, doi:10.1088/1126-6708/2008/12/039.

[64] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, J. Winter, Event generation with SHERPA 1.1, JHEP 02 (2009) 007. arXiv:0811.4622.

[65] E. Bothmann, et al., Event Generation with Sherpa 2.2, SciPost Phys. 7 (3) (2019) 034. arXiv:1905.09127.

[66] A. Gehrmann-De Ridder, T. Gehrmann, G. Heinrich, Four particle phase space integrals in massless QCD, Nucl. Phys. B 682 (2004) 265–288. arXiv:hep-ph/0311276.

[67] J. J. Lopez-Villarejo, P. Z. Skands, Efficient Matrix-Element Matching with Sector Showers, JHEP 11 (2011) 150. arXiv:1109.3608.

[68] S. Höche, F. Krauss, M. Schönherr, F. Siegert, Automating the POWHEG method in Sherpa, JHEP 04 (2011) 024. arXiv:1008.5399, doi:10.1007/JHEP04(2011)024.

[69] W. T. Giele, G. C. Stavenga, J.-C. Winter, A Forward Branching Phase-Space GeneratorarXiv:1106.5045.

[70] W. T. Giele, Improved Partonic Event Generators at Lepton CollidersarXiv:1504.02137.

[71] T. M. Figy, W. T. Giele, A Forward Branching Phase Space Generator for Hadron colliders, JHEP 10 (2018) 203. arXiv:1806.09678.
[72] T. Chen, T. M. Figy, W. T. Giele, A Projective Phase Space Generator for Hadronic Vector Boson Plus One Jet Production (2019). arXiv:1907.03893.
[73] J. M. Campbell, W. T. Giele, C. Williams, The Matrix Element Method at Next-to-Leading Order, JHEP 11 (2012) 043. arXiv:1204.4424.
[74] S. Weinzierl, NNLO corrections to 2-jet observables in electron-positron annihilation, Phys. Rev. D 74 (2006) 014020. arXiv:hep-ph/0606008.
[75] K. G. Chetyrkin, J. H. Kuhn, A. Kwiatkowski, QCD corrections to the e^+e^- cross-section and the Z boson decay rate, Phys. Rept. 277 (1996) 189–281. arXiv:hep-ph/9503396.
[76] S. G. Gorishnii, A. L. Kataev, S. A. Larin, L. R. Surguladze, Corrected Three Loop QCD Correction to the Correlator of the Quark Scalar Currents and Γ(Tot) ($H^0 \to$ Hadrons), Mod. Phys. Lett. A 5 (1990) 2703–2712.
[77] K. G. Chetyrkin, Correlator of the quark scalar currents and Gamma(tot) (H —> hadrons) at O (alpha-s**3) in pQCD, Phys. Lett. B 390 (1997) 309–317. arXiv:hep-ph/9608318.
[78] P. A. Baikov, K. G. Chetyrkin, J. H. Kuhn, Scalar correlator at O(alpha(s)**4), Higgs decay into b-quarks and bounds on the light quark masses, Phys. Rev. Lett. 96 (2006) 012003. arXiv:hep-ph/0511063.
[79] V. Del Duca, C. Duhr, G. Somogyi, F. Tramontano, Z. Trócsányi, Higgs boson decay into b-quarks at NNLO accuracy, JHEP 04 (2015) 036. arXiv:1501.07226.
[80] K. G. Chetyrkin, B. A. Kniehl, M. Steinhauser, Hadronic Higgs decay to order alpha-s**4, Phys. Rev. Lett. 79 (1997) 353–356. arXiv:hep-ph/9705240.

8

QCD Radiation in VBF Higgs Production

The precise determination of the properties of the Higgs boson offers a possible guide to physics beyond the Standard Model. A peculiarly interesting channel is the production of Higgs bosons in vector boson fusion (VBF). At LO, a VBF process consists of two quarks that are forward-scattered by emitting vector bosons, which in turn fuse to a colour-singlet particle such as a Z^0 or H^0 boson, *e.g.* $qq' \to qq'H^0$. A closely related class of processes is vector boson scattering (VBS), in which the t-channel vector bosons scatter to produce two outgoing vector bosons, *e.g.* $qq' \to qq'W^+W^-$ Due to its distinct colour topology, given by two disconnected initial-final colour flows at LO, a very perceptible event signature is inherent to VBF processes. While the colour-singlet particle is located in the central rapidity region, QCD radiation off the initial-final quark pairs is situated further towards larger rapidities, with a strong suppression at central rapidities due to destructive interference in coherent radiation off the quark lines. This distinct event topology allows to employ a specific set of fiducial cuts in experimental analyses usually known as VBF cuts: for any event to be considered a VBF event, two hard jets with a large collective invariant mass and large rapidity separation must be present in opposite detector hemispheres. These two hard jets are then called tagging jets and can be used to substantially suppress background processes such as gluon fusion with two additional jets. The VBF channel therefore delivers a clean environment for precision studies pertaining to the nature of the Higgs boson.

In spite of the clean experimental environment, the VBF topology poses a challenge for parton showers, as the coherent suppression of QCD radiation in the central rapidity region must be faithfully modelled. For instance, it has been known for some time that Pythia's default transverse-momentum-ordered shower fails to reproduce the correct radiation pattern in such topologies [431, 432] due to its use of a global recoil scheme for initial-state branchings in conjunction with DGLAP splitting functions, which by themselves are incapable of describing coherent soft radiation. It was demonstrated that a better description can be obtained with Pythia's non-default dipole-recoil option, which replaces the independent DGLAP evolution of the initial-state and final-state parton in such initial-final colour-dipoles by a coherent dipole/antenna-like branching of both dipole ends. Due to the importance for future electroweak and Higgs precision measurements, a lot of effort has concentrated on the determination of systematic uncertainties pertaining to the choice of shower models and NLO matching schemes in VBF and VBS processes, *cf. e.g.* [433–436].

As VINCIA's (sector) antenna showers are anchored in the soft limit, it can be assumed that VINCIA provides a more meaningful description of QCD radiation in VBF processes than PYTHIA's default shower, comparable to that of PYTHIA's dipole-recoil shower. Section 8.1 contains a study of QCD radiation in VBF Higgs production, taking VINCIA as a baseline and comparing it to both PYTHIA's default and dipole-recoil shower options. A set of observables deemed especially sensitive to the radiation pattern in VBF processes is considered and the LO predictions using the three showers are confronted with NLO-matched results in the POWHEG scheme using POWHEGBOX [229, 437]. To this end, dedicated `PowhegHooks` are introduced for VINCIA. Furthermore, the effect of including up to four additional hard jets via VINCIA's CKKW-L implementation, *cf.* chapter 5, is studied using SHERPA to generate the tree-level event samples. It should be pointed out that currently VINCIA is the only shower in PYTHIA 8.3 that supports merging in VBF/VBS processes. The application to a non-trivial process also highlights the operability of the VINCIA shower model in PYTHIA 8.3 in practice.

8.1 Published Material

A Study of QCD Radiation in VBF Higgs Production with VINCIA and PYTHIA

Stefan Höche, Stephen Mrenna, Shay Payne, **Christian T Preuss**, and Peter Skands

Submitted to SciPost Physics
e-Print: arXiv:2106.10987 [hep-ph]

Since the submission of this thesis, a revised version of the paper has been accepted for publication by SciPost Physics.

Paper begins overleaf.

FERMILAB-PUB-21-289-SCD-T, MCNET-21-11

A Study of QCD Radiation in VBF Higgs Production with Vincia and Pythia

Stefan Höche[1], Stephen Mrenna[1], Shay Payne[2], Christian T Preuss[2*], Peter Skands[2]

1 Fermi National Accelerator Laboratory, Batavia, IL, 60510, USA
2 School of Physics and Astronomy, Monash University, Wellington Road, Clayton, VIC-3800, Australia
* christian.preuss@monash.edu

August 19, 2021

Abstract

We discuss and illustrate the properties of several parton-shower models available in Pythia and Vincia, in the context of Higgs production via vector boson fusion (VBF). In particular, the distinctive colour topology of VBF processes allows to define observables sensitive to the coherent radiation pattern of additional jets. We study a set of such observables, using the Vincia sector-antenna shower as our main reference, and contrast it to Pythia's transverse-momentum-ordered DGLAP shower as well as Pythia's dipole-improved shower. We then investigate the robustness of these predictions as successive levels of higher-order perturbative matrix elements are incorporated, including next-to-leading-order matched and tree-level merged calculations, using Powheg Box and Sherpa respectively to generate the hard events.

Contents

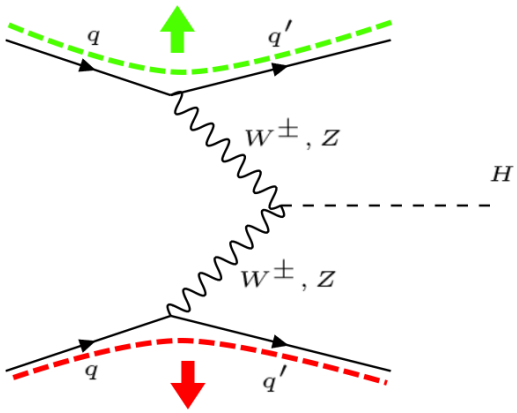


Figure 1: QCD colour flow of the LO VBF Higgs production process. Due to the kinematics of the interaction, QCD radiation is directed in the forward region of the detector.

1 Introduction

Higgs boson production via Vector Boson Fusion (VBF) — fig. 1 — is among the most important channels for Higgs studies at the Large Hadron Collider (LHC). With a Standard-Model (SM) cross section of a few pb at LHC energies, VBF accounts for order 10% of the total LHC Higgs production rate [1]. The modest rate is compensated for by the signature feature of VBF processes: two highly energetic jets generated by the scattered quarks, in the forward and backward regions of the detector respectively, which can be tagged experimentally and used to significantly reduce background rates. Moreover, the distinct colour flow of the VBF process at leading order (LO), highlighted by the coloured thick dashed lines in fig. 1, strongly suppresses any coherent bremsstrahlung into the central region, leaving this region comparatively clean and well suited for precision studies of the Higgs boson decay products. With over half a million Higgs bosons produced in the VBF channel in total during Run II of the LHC and a projection that this will more than double during Run III, studies of this process have already well and truly entered the realm of precision physics.

On the theory side, the current state of the art for the $\mathrm{H} + 2j$ process in fixed-order perturbation theory is inclusive next-to-next-to-next-to-leading order QCD [2], fully differential next-to-next-to-leading order (NNLO) QCD [3–6] and next-to-leading-order (NLO) elec-

troweak (EW) calculations [7]. These calculations of course only offer their full precision for observables that are non-zero already at the Born level, such as the total cross section and differential distributions of the Higgs boson and tagging jets. For more exclusive event properties, such as bremsstrahlung and hadronisation corrections, the most detailed description is offered by combinations of fixed-order and parton-shower calculations. To this end, two recent phenomelogical studies [8, 9] compared different NLO+PS simulations among each other as well as to NLO and NNLO calculations. These comparative studies catered to two needs; firstly, the reliability of matched calculations was tested in regions where resummation effects are small. Furthermore, a more realistic estimate of parton-shower as well as matching uncertainties was obtained by means of different shower and matching methods in independent implementations.

The earlier of the two studies [8] highlighted that different NLO+PS implementations describe the intrinsically coherent radiation in this process quite differently, and that the uncertainties arising from the choice of the shower and matching implementation can persist even at the NLO-matched level. Among its central results, the study [8] confirmed the observation of [10] that Pythia's default shower [11–13] describes the emission pattern of the third jet poorly, essentially missing the coherence of the initial-final dipoles. This effect was most pronounced for MadGraph_aMC@Nlo [14] + Pythia, for which a global recoil scheme must be used in both the time-like and space-like shower in order to match the subtraction terms implemented in MadGraph_aMC@Nlo. For Powheg-Box [15] + Pythia, the difference persisted when using the global recoil scheme[1]. However, changing to Pythia's alternative dipole-recoil scheme [16], which should reproduce coherence effects more faithfully, improved the agreement, both with calculations starting from H$+3j$ as well as with the angular-ordered coherent shower model in Herwig 7 [17].

The more recent study [9] highlighted a number of interesting aspects of vector boson fusion that can be exploited to enhance the signal-to-background ratio in future measurements: Firstly, if the Higgs boson is boosted, the t-channel structure of the VBF matrix elements leads to less QCD radiation when compared to the irreducible background from gluon-gluon fusion. Secondly, it was found that a global jet veto provides a similarly effective cut as a central jet veto, leading to much reduced theoretical uncertainties, and in particular eliminating the need to resum non-global logarithms associated with inhibited radiation in the rapidity gap. Despite a good overall agreement between fixed-order NNLO and NLO-matched parton shower predictions, the study also pointed out a few subtle disagreements for highly boosted Higgs boson topologies. In these scenarios, the standard fixed-order paradigm of operating with a single factorisation scale is no longer appropriate, because higher-order corrections should be resummed individually for the two impact factors in the structure-function approach.

The uncertainties arising from matching systematics in vector-boson-fusion and vector-boson-scattering processes (VBS) have also been studied in the past [18] with rather good agreement between different showers at the level of H$+3j$ NLO+PS calculations [19], although in that study, only the Powheg matching scheme was considered. Very recently, two extensive reviews [20, 21] collected experimental results and theoretical developments in VBS processes in view of the high-luminosity upgrade of the LHC as well as future colliders. A summary of Monte Carlo event generators used in the modelling of VBS processes in ATLAS was presented in [22].

[1] We note that the global recoil scheme is the default choice only for Pythia's space-like DGLAP shower, while the time-like DGLAP shower uses a dipole-like recoil scheme per default.

On the experimental side, recent studies of VBF Higgs production by ATLAS [23,24] and CMS [25,26] have used Pythia's default shower model matched to the NLO via the Powheg technique, with only one of them [23] employing Pythia's dipole option. The associated modelling uncertainties, and ways to reduce them, therefore remain of high current relevance.

We extend the comparative study of [8] to include the new Vincia sector-antenna shower [27] that has become available starting from Pythia version 8.304. Based on findings pertaining to antenna [28–31] and dipole [32–35] showers, we expect that, at least at leading colour, Vincia's showers capture QCD coherence effects in VBF more accurately than Pythia's default shower. To this end, we note that the emitter-recoiler agnostic antenna recoil employed in Vincia is free of adverse kinematic effects [36]. We also consider two new observables designed to further probe the amount of coherent radiation by measuring the summed transverse energy H_T for $|\eta| < 0.5$ and for $|\eta - \eta_0| < 0.5$ respectively, where η_0 is the midpoint between the two tagging jets. To investigate the robustness of the predictions, we include not only Powheg-Box + Pythia [13, 15] but also a new dedicated implementation of the CKKW-L merging scheme [37–39] for sector showers [40], with hard events with up to four additional jets generated by Sherpa 2 [41, 42]. We emphasise that this is currently the only multi-jet merging approach in Pythia 8.3 which can handle VBF processes[2]. Additionally, we highlight the systematic uncertainties arising from the use of vetoed showers in the Powheg scheme and make recommendations for settings related to the use of these in Pythia.

This study is structured as follows. We begin with an overview of the setup for our simulations in section 2; starting with an overview of the fixed order, shower, matched and merged calculations and leading towards a description of the analysis we perform. We then move on to discuss the results of our analysis in section 3, with our conclusions and recommendations listed in section 4.

2 Setup of the Simulation

We consider Higgs production via VBF in proton-proton collisions at the high-luminosity LHC with a centre-of-mass energy of $\sqrt{s} = 14$ TeV.

The simulation is factorised into the generation of the hard process using Sherpa 2 (for the LO merging samples) and Powheg-Box v2 (for the NLO matched samples) and subsequent showering with Pythia 8.306. A cross check is also performed using Pythia's internal Born-level VBF process. Details on the hard-process setups are given in section 2.1.

Since we expect the Vincia antenna shower to account for coherence more faithfully than does Pythia's default "simple" $p_\perp$-ordered DGLAP shower, we take Vincia's description as the baseline for our comparisons, contrasting it to Pythia's default and "dipole-recoil" options. Details on the shower setups are given in section 2.2.

Higher fixed-order corrections are taken into account at NLO+PS accuracy via the Powheg scheme, and for Vincia also in the CKKW-L scheme up to $\mathcal{O}(\alpha_S^4)$. We expect that these corrections will be smaller for coherent shower models than for incoherent ones, hence these comparisons serve both to test the reliability of the baseline showers and to illustrate any ambiguities that remain after these corrections are included. Details on the matching and merging setups are given in section 2.3.

[2] We do note that a technical (but due to the use of incoherent IF kinematics unphysical) fix was introduced in Pythia 8.242 and is planned to be re-implemented in a future version of Pythia 8.3.

Finally, in section 2.4, we define the observables and the VBF analysis cuts that are used for the numerical studies in section 3.

Note that, since we are primarily interested in exploring the coherence properties of the perturbative stages of the event simulation, most of the results will be at the so-called "parton level", i.e. without accounting for non-perturbative or non-factorisable effects such as hadronisation, primordial k_T, or multi-parton interactions (MPI). Although this is not directly comparable to physical measurements (nor is the definition universal since different shower models define the cutoff differently), the factorised nature of the infrared and collinear safe observables we consider imply that, while non-perturbative effects may act to smear out the perturbative differences and uncertainties, they would not in general be able to obviate them, thus making studies of the perturbative stages interesting in their own right. Nevertheless, with jet p_T values going down to 25 GeV and H_T being sensitive to the overall amount of energy scattered into the central region, we include further comparisons illustrating the effect of non-perturbative corrections at the end of section 3.

2.1 Hard Process

For the parton-level event generation, we use a stable Higgs boson with a mass of $M_H = 125$ GeV, and we set the electroweak boson masses and widths to

$$\begin{aligned} M_Z &= 91.1876 \text{ G}e\text{V}\,, \quad \Gamma_Z = 2.4952 \text{ G}e\text{V}\,, \\ M_W &= 80.385 \text{ G}e\text{V}\,, \quad \Gamma_W = 2.085 \text{ G}e\text{V}\,. \end{aligned} \tag{1}$$

Electroweak parameters are derived from this set with the additional input of the electromagnetic coupling constant at the Z pole ($\alpha(M_Z)$ scheme, `EW_SCHEME = 2` in Sherpa):

$$\frac{1}{\alpha(M_Z)} = 128.802\,. \tag{2}$$

We treat all flavours including the bottom quark as massless and use a diagonal CKM mixing matrix. In both Sherpa and Powheg-Box, we use the `CT14_NNLO_as118` [43] PDF set provided by Lhapdf6 [44] with the corresponding value of α_S. For the sample generated with Pythia's internal VBF implementation, we use its default `NNPDF23_lo_as_0130_qed` PDF set [45, 46].

We consider only VBF topologies, neglecting Higgsstrahlung contributions which appear at the same order in the strong and electroweak coupling. Identical-flavour interference effects are neglected in events generated with Powheg-Box and Pythia, but are included in events obtained with Sherpa, although their impact was found to be small [9]. At NLO, the process is calculated in the structure function approximation, neglecting interferences between the two quark lines. For both, internal and external events, only a single scale will be assigned per event, notwithstanding that different scales could in principle be assigned to the two forward-scattered quarks. Differences pertaining to the scale assignment in internal and external events will be discussed in section 3.1.

Tree-level event samples with up to four additional jets are generated using an HPC-enabled variant of Sherpa 2 [41, 42], utilising the Comix matrix-element generator [47]. To facilitate efficient parallelised event generation and further processing, events are stored in the binary HDF5 data format [42]. The factorisation and renormalisation scales are chosen

to be

$$\mu_{\mathrm{F}}^2 = \mu_{\mathrm{R}}^2 = \frac{\hat{H}_{\mathrm{T}}^2}{4} \quad \text{with} \quad \hat{H}_{\mathrm{T}} = \sum_j p_{\mathrm{T},j} + \sqrt{M_{\mathrm{H}}^2 + p_{\mathrm{T,H}}^2}\,. \tag{3}$$

and jets are defined according to the k_{T} clustering algorithm with $R = 0.4$ and a cut at 20 GeV.

Pythia's internal events are generated with scales governed by the two switches `SigmaProcess:factorScale3VV` and `SigmaProcess:renormScale3VV`, respectively. Their default values `= 2` and `= 3`, respectively, correspond to the choices

$$\mu_{\mathrm{F}}^2 = \sqrt{m_{\mathrm{T,V_1}}^2 m_{\mathrm{T,V_2}}^2} \equiv \sqrt{(M_{\mathrm{V_1}}^2 + p_{\mathrm{T,q_1}}^2)(M_{\mathrm{V_2}}^2 + p_{\mathrm{T,q_2}}^2)}\,, \tag{4}$$

$$\mu_{\mathrm{R}}^2 = \sqrt{m_{\mathrm{T,V_1}}^2 m_{\mathrm{T,V_2}}^2 m_{\mathrm{T,H}}^2} \equiv \sqrt[3]{(M_{V_1}^2 + p_{\mathrm{T,q_1}}^2)(M_{\mathrm{V_2}}^2 + p_{\mathrm{T,q_2}}^2) m_{\mathrm{T,H}}^2}\,, \tag{5}$$

with the pole masses of the exchanged vector bosons $M_{\mathrm{V_1}}$, $M_{\mathrm{V_2}}$, the transverse mass of the Higgs boson $m_{\mathrm{T},H}$, and the transverse momenta of the two final-state quarks p_{T,q_1}, p_{T,q_2}.

For NLO calculations matched to parton showers, we consider the Powheg [48, 49] formalism. Powheg samples are generated with Powheg-Box v2 [15, 50] with the factorisation and renormalisation scales chosen as

$$\mu_{\mathrm{F}}^2 = \mu_{\mathrm{R}}^2 = \frac{M_{\mathrm{H}}}{2}\sqrt{\left(\frac{M_{\mathrm{H}}}{2}\right)^2 + p_{\mathrm{T,H}}^2}\,. \tag{6}$$

Since the study in [8] did not find any significant effect from the choice of the "`hdamp`" parameter in Powheg, we do not include any such damping here, corresponding to a choice of `hdamp = 1`.

2.2 Showers

The hard events defined above are showered with the three following shower models, which are all available in Pythia 8.306:

- Vincia's sector antenna shower [27]. The "sector" mode is the default option for Vincia since Pythia 8.304 and also enables us to make use of Vincia's efficient CKKW-L merging [40]. We expect it to exhibit the same level of coherence as the fixed-order matrix elements, at least at leading colour (LC), since its QCD antenna functions and corresponding phase-space factorisations explicitly incorporate the soft-eikonal function for all possible (LC) colour flows. Of particular relevance to this study is its coherent treatment of "initial-final" (IF) colour flows.

- Pythia's default "simple shower" model [11, 12], which implements $p_\perp$-ordered DGLAP evolution with dipole-style kinematics. For IF colour flows, however, the kinematic dipoles are not identical to the colour dipoles, and this can impact coherence-sensitive observables [51].

- Pythia's "simple shower" with the dipole-recoil option [16]. Despite its name, this not only changes the recoil scheme; in fact, it replaces the two independent DGLAP evolutions of IF dipoles by a coherent, antenna-like, dipole evolution, while keeping the DGLAP evolution of other dipoles unchanged. This option should therefore lead to radiation patterns exhibiting a similar level of coherence as Vincia.

Ordinarily, PYTHIA would of course also add decays of the Higgs boson, and any final-state radiation associated with that. However, as a colour-singlet scalar with $\Gamma_H \ll \Lambda_{QCD}$ and $\Gamma_H/M_H \sim \mathcal{O}(10^{-5})$, its decay can be treated as factorised from the production process to a truly excellent approximation. For the purpose of this study, we therefore keep the Higgs boson stable, to be able to focus solely on the radiation patterns of the VBF production process itself, without the complication of decay products in the central region.

For all of the shower models, we retain PYTHIA's default PDF choice[3], regardless of which PDF set was used to generate the hard process. This is done to remain consistent with the default shower tunings [52] and due to the better-controlled backwards-evolution properties of the default set [53].

Per default, the shower starting scale is chosen to be the factorisation scale of the hard process,

$$\mu^2_{\mathrm{PS}} = \mu^2_{\mathrm{F}}\,. \tag{7}$$

In VINCIA, this scale can be varied by a multiplicative "fudge" factor, controlled by `Vincia:pTmaxFudge`,

$$\mu^2_{\mathrm{PS}} = k_{\mathrm{fudge}}\,\mu^2_{\mathrm{F}}\,,$$

while in PYTHIA, the starting scales of the initial-state and final-state showers can be varied independently,

$$\mu^2_{\mathrm{PS,FSR}} = k_{\mathrm{fudge,FSR}}\,\mu^2_{\mathrm{F}}\,,$$
$$\mu^2_{\mathrm{PS,ISR}} = k_{\mathrm{fudge,ISR}}\,\mu^2_{\mathrm{F}}\,,$$

controlled by `TimeShower:pTmaxFudge` and `SpaceShower:pTmaxFudge`, respectively.

In a similar vein, the strong coupling in the shower is evaluated at the shower p_{T}-scale[4], modified by renormalisation-scale factors k_{ren}. In PYTHIA, the strong coupling at the Z mass is set to $\alpha_{\mathrm{S}}(M_{\mathrm{Z}}) = 0.1365$ and independent scale factors for ISR and FSR are implemented,

$$\alpha_{\mathrm{S}}^{\mathrm{Pythia,FSR}}(p^2_{\perp\mathrm{evol,FSR}}) = \alpha_{\mathrm{S}}^{\overline{\mathrm{MS}}}(k_{\mathrm{R,FSR}}\,p^2_{\perp\mathrm{evol,FSR}})\,,$$
$$\alpha_{\mathrm{S}}^{\mathrm{Pythia,ISR}}(p^2_{\perp\mathrm{evol,ISR}}) = \alpha_{\mathrm{S}}^{\overline{\mathrm{MS}}}(k_{\mathrm{R,ISR}}\,p^2_{\perp\mathrm{evol,ISR}})\,.$$

These can be set via `TimeShower:renormMultFac` and `SpaceShower:renormMultFac`, respectively, and are unity by default. The transverse-momentum evolution variables $p^2_{\perp\mathrm{evol,FSR}}$ and $p^2_{\perp\mathrm{evol,ISR}}$ are defined as in [11].

For VINCIA, on the other hand, a more refined choice can be made with separate renormalisation factors being implemented for (initial- and final-state) emissions, (initial- and final-state) gluon splittings, and (initial-state) quark conversions. These have the default settings:

$$k^{\mathrm{F}}_{\mathrm{R,Emit}} = 0.66\,, \quad k^{\mathrm{F}}_{\mathrm{R,Split}} = 0.8\,,$$
$$k^{\mathrm{I}}_{\mathrm{R,Emit}} = 0.66\,, \quad k^{\mathrm{I}}_{\mathrm{R,Split}} = 0.5\,, \quad k^{\mathrm{I}}_{\mathrm{R,Conv}} = 0.5\,,$$

which can be set via the parameters

`Vincia:renormMultFacEmitF`
`Vincia:renormMultFacSplitF`
`Vincia:renormMultFacEmitI`
`Vincia:renormMultFacSplitI`
`Vincia:renormMultFacConvI`.

[3] `NNPDF23_lo_as_0130_qed`.

[4] We refer to the argument of the strong coupling used in the shower as the shower renormalisation scale.

Additionally, VINCIA uses the CMW scheme [54] (while PYTHIA does not), i.e. it evaluates the strong coupling according to

$$\alpha_{\mathrm{S}}^{\mathrm{CMW}} = \alpha_{\mathrm{S}}^{\overline{\mathrm{MS}}} \left(1 + \frac{\alpha_{\mathrm{S}}^{\overline{\mathrm{MS}}}}{2\pi} \left[C_A \left(\frac{67}{18} - \frac{\pi^2}{6} \right) - \frac{5 n_f}{9} \right] \right) , \tag{8}$$

where $\alpha_{\mathrm{S}}^{\overline{\mathrm{MS}}}(M_{\mathrm{Z}}) = 0.118$, so that

$$\alpha_{\mathrm{S}}^{\mathrm{Vincia}}(p_\perp^2) = \alpha_{\mathrm{S}}^{\mathrm{CMW}}(k_{\mathrm{R}}\, p_\perp^2) \tag{9}$$

with the VINCIA evolution variable as defined in [27].

2.3 Matching and Merging

In the following, we will briefly review the defining features of the POWHEG NLO matching and the CKKW-L merging schemes we will use in this study. In particular, we will focus on the technicalities and practicalities to ensure a consistent use. Detailed reviews of the POWHEG schemes can for instance be found in [55] and [56]. The CKKW-L scheme is explained in detail in [39] and its extension to the VINCIA sector shower in [40].

2.3.1 POWHEG Matching

In the POWHEG formalism, events are generated according to the inclusive NLO cross section with the first emission generated according to a matrix-element corrected no-emission probability.

Since the shower kernels in the POWHEG no-emission probability are replaced by the ratio of the real-radiation matrix element to the Born-level one, it is independent of the shower it will later be matched to. It is, however, important to stress that generally, the POWHEG ordering variable will not coincide with the ordering variable of the shower. Starting a shower with a different ordering variable at the POWHEG scale of the first emission might thus lead to over- or undercounting emissions. A simple method to circumvent this was presented in [57]. There, the shower is started at the phase space maximum (a so-called"power shower" [58]) and emissions harder than the POWHEG one are vetoed until the shower reaches a scale below the scale of the first emission. For general ordering variables, there is, however, no guarantee that once the shower falls below the scale of the POWHEG emission it will not generate a harder emission later on in the evolution. This is especially important if the shower is not ordered in a measure of hardness but e.g. in emission angles, such as the HERWIG $\tilde{q}$ shower [59]. In these cases, it is advisable to recluster the POWHEG emission and start a truncated and vetoed shower off the Born state [48], see also [60–62] for the use of truncated showers in merging schemes. This scheme also avoids the issue that in vetoed showers, all emissions in the shower off a Born+1-jet state are compared against the POWHEG emission as if they were the first emission themselves. But from the point of view of kinematics and colour they will still be the second, third, etc.

However, since all showers we consider here are ordered in a notion of transverse momentum, it shall suffice for our purposes to use the simpler "vetoed power shower" scheme. To this end, we have amended the existing POWHEG user hook for PYTHIA's showers by a dedicated one for POWHEG+VINCIA, which has been included in the standard release of PYTHIA starting from version 8.306; see appendix A for detailed instructions.

For both PYTHIA and VINCIA, we use a vetoed shower with the POWHEG p_{T} and d_{ij} definitions, corresponding to the mode `POWHEG:pTdef = 1`. We define the POWHEG scale with respect to the radiating leg and use PYTHIA's definition of emitter and recoiler, corresponding to the modes `POWHEG:pTemt = 0` and `POWHEG:emitted = 0`. Per default, we choose to define the scale of the POWHEG emission by the minimum p_{T} among all final-state particles, i.e. use `POWHEG:pThard = 2`, according to the suggestion in [63]. As an estimate of the uncertainty of this choice, we vary the $p_{\mathrm{T,hard}}$ scale to be the LHEF scale and the p_{T} of the POWHEG emission, corresponding to the modes `POWHEG:pThard = 0` and `POWHEG:pThard = 1`, respectively.

The purpose of these settings is to ensure maximally consistent scale definitions while not reverting to the (more involved) "truncated and vetoed shower" scheme mentioned above. While we deem the choices made here appropriate for the case at hand they remain ambiguous, effectively introducing systematic matching uncertainties into the (precision) calculation. As a means of estimating these uncertainties, we will discuss the influence of the $p_{\mathrm{T,hard}}$ scale setting on physical observables below in section 3.

2.3.2 CKKW-L Merging

Multi-leg merging schemes aim at correcting parton shower predictions away from the soft and collinear regions. In the CKKW-L merging scheme [39], multiple inclusive tree-level event samples are combined to a single inclusive one by introducing a (somewhat arbitrary) "merging scale" t_{MS} which separates the matrix-element ($t > t_{\mathrm{MS}}$) from the parton-shower ($t < t_{\mathrm{MS}}$) region. In this way, over-counting of emissions is avoided while accurate parton-shower resummation in logarithmically enhanced regions and leading-order accuracy in the regions of hard, well-separated jets is ensured if the merging scale is chosen appropriately.

The missing Sudakov suppression in higher-multiplicity configurations is calculated post-facto by the use of truncated trial showers between the nodes of the most probable "shower history". In this context, the shower history represents the sequence of intermediate states the parton shower at hand would (most probably) have generated to arrive at the given n-jet state. Usually, this sequence is constructed by first finding all possible shower histories and subsequently choosing the one that maximises the branching probability, i.e., the product of branching kernels and the Born matrix element. As we employ this scheme with VINCIA's sector shower, a few comments are in order. The objective of the sector shower is to replace the probabilistic shower history by a deterministic history, governed by the singularity structure of the matrix element. This means that at each point in phase space only the most singular branching contributes. In the shower, this is ensured by vetoing any branchings that do not abide by this; in the merging, this results in a faster and less resource-intensive algorithm, as it is no longer required to generate a large number of possible histories. Details and subtleties of VINCIA's sectorised CKKW-L implementation can be found in [40].

The CKKW-L merging scheme is in principle implemented for all showers in PYTHIA 8.3. However, the intricate event topology of VBF processes currently prohibits the use of PYTHIA's default merging implementation[5]. We hence limit ourselves to study the effect of merging with VINCIA, and have adapted VINCIA's CKKW-L implementation [40] so that VBF processes are consistently treated. Specifically, the flag `Vincia:MergeVBF = on` should be used, which restricts the merging to only consider shower histories that retain exactly two initial-final quark lines. As a consequence, there must not be any "incomplete histories"

[5] We note that a technical fix for this was available in PYTHIA 8.245 and will become available again in PYTHIA 8.3 in the future.

(histories that do not cluster back to a VBF Born configuration); this should be guaranteed as long as the input event samples are of the VBF type only and no QED or EW emissions are generated. A complete list of relevant settings for the use of VINCIA's CKKW-L merging is collected in appendix B.

2.4 Analysis

We use the anti-k_{T} algorithm [64] with $R = 0.4$, as implemented in the FASTJET [65] package, to cluster jets in the range,

$$p_{\mathrm{T}} > 25 \text{ GeV}\,, \quad |\eta| < 4.5\,.$$

In addition, we employ typical VBF cuts to ensure that the two "tagging jets" are sufficiently hard, have a large separation in pseudorapidity, and are located in opposite hemispheres:

$$m_{j_1,j_2} \geq 600 \text{ GeV}\,, \quad |\Delta\eta_{j_1,j_2}| \geq 4.5\,, \quad \eta_{j_1} \cdot \eta_{j_2} \leq 0\ .$$

We consider the following observables:

- **Pseudorapidity Distributions:** at the Born level, the two tagging jets already have nontrivial pseudorapidity distributions. These are sensitive to showering chiefly via recoil effects and via the enhancement of radiation towards the beam directions. The third (and subsequent) jets are of course directly sensitive to the generated emission spectra. To minimise contamination from final-state radiation off the tagging jets, we also consider the pseudorapidity of the radiated jet(s) relative to the midpoint of the two tagging jets,

  $$\eta^*_{j_i} = \eta_{j_i} - \eta_0\ , \tag{10}$$

  with the midpoint defined by:

  $$\eta_0 = \tfrac{1}{2}(\eta_{j_1} + \eta_{j_2})\ . \tag{11}$$

- **Transverse Momentum Distributions:** we expect coherence effects for the radiated jets ($i > 2$) to be particularly pronounced for radiation that is relatively soft in comparison to the characteristic scale of the hard process. Conversely, the transverse momenta of the two tagging jets should mainly be affected indirectly, via momentum-conservation (recoil) effects.

- **Scalar Transverse Momentum Sum:** as a more inclusive measure of the summed jet activity in the central rapidity region, we consider the scalar transverse momentum sum of all reconstructed jets (defined as above, i.e., with $p_{\mathrm{T}} > 25$ GeV),

  $$H_{\mathrm{T}} = \sum_j |p_{\mathrm{T},j}|\,, \tag{12}$$

  in two particular regions:

  - in the central rapidity region, $\eta \in \left[-\frac{1}{2}, +\frac{1}{2}\right]$
  - around the midpoint of the tagging jets, $\eta^* \in \left[-\frac{1}{2}, +\frac{1}{2}\right]$, cf eq. (10).

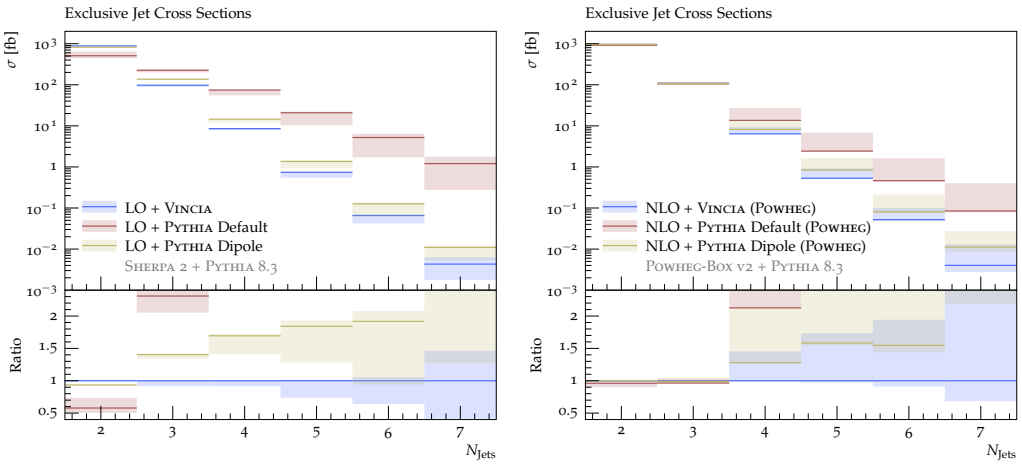


Figure 2: Exclusive jet cross sections at LO+PS (*left*) and POWHEG NLO+PS (*right*) accuracy. The bands are obtained by a variation of the default shower starting scale by a factor of two or the variation of the hard scale, respectively.

We point out that, due to the way it is constructed, the second of these regions is not sensitive to the tagging jets, as it is not possible for them to fall within this region. Unlike the previous two observables, H_{T} is sensitive to the overall radiation effect in the given region, not just that of a certain jet multiplicity. As such, we expect H_{T} to give a measure of the all-orders radiation effects.

The analysis is performed using the RIVET analysis framework [66, 67] and based on the one used in [8].

3 Results

In this section, we present the main results of our study based on the setup described in the last section. In fig. 2, the exclusive jet cross sections for up to 7 jets are shown at LO+PS and NLO+PS (via the POWHEG scheme) accuracy at the Born level. While there are very large differences between the three shower predictions at the leading order, there is good agreement between the NLO+PS predictions at least for the 2- and 3-jet cross sections.

3.1 Leading Order

It is instructive to start by studying the properties of the baseline leading-order + shower calculations, without including higher fixed-order corrections.

We use the leading-order event samples generated with SHERPA and by default let the factorisation scale μ_{F}^2 define the shower starting scale. As a way to estimate the uncertainty associated with this choice, we vary the shower starting scale μ_{PS}^2 by a factor $k_{\mathrm{fudge}} \in \left[\frac{1}{2}, 2\right]$, $\mu_{\mathrm{PS}}^2 = k_{\mathrm{fudge}}\mu_{\mathrm{F}}^2$. Strictly speaking, shower starting scales not equal to the factorisation scale lead to additional PDF ratios in the no-branching probabilities generated by the shower, but for factor-2 variations these are consistent with unity (since the PDF evolution is logarithmic) and we therefore neglect them. Compared to the shower starting scale, variations of the shower

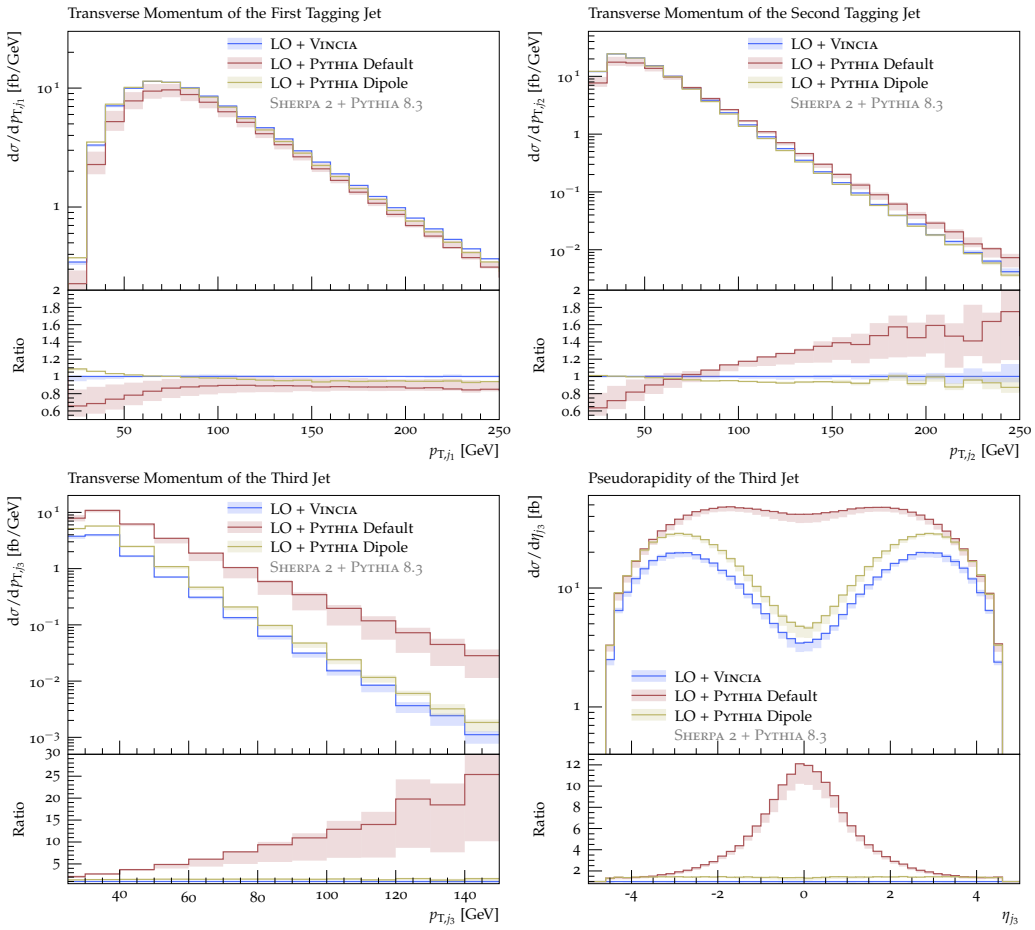


Figure 3: Transverse momentum of the first tagging jet (*top left*), second tagging jet (*top right*), third jet (*bottom left*), and pseudorapidity of the third jet (*bottom right*) at LO+PS accuracy. The bands are obtained by a variation of the default shower starting scale by a factor of two.

renormalisation scale only have a marginal effect and are therefore not shown here. As we are primarily concerned with the shower radiation patterns, we do not vary the scales in the fixed-order calculation. The effect of those variations have been studied extensively in the literature before, cf. e.g. [8, 18].

In fig. 3, the transverse momentum distributions of the two tagging jets and as well as the transverse momentum and pseudorapidity distributions of the third-hardest jet are shown. While the tagging jet p_{T} spectra agree well between Vincia and Pythia with dipole recoil, differences are visible for the third-jet observables, with similar shapes but a slightly larger rate produced by the Pythia dipole-recoil shower. The distributions obtained with the Pythia default shower, on the other hand, neither agree in shape nor in the rate with the other two. In fact, almost no suppression of radiation in the central-rapidity region is visible and the shower radiation appears at a much higher transverse momentum scale. The high emission rate in the default shower also implies that the tagging jets receive much larger corrections with this shower than with the other models, as evident from the tagging jet p_{T} distributions.

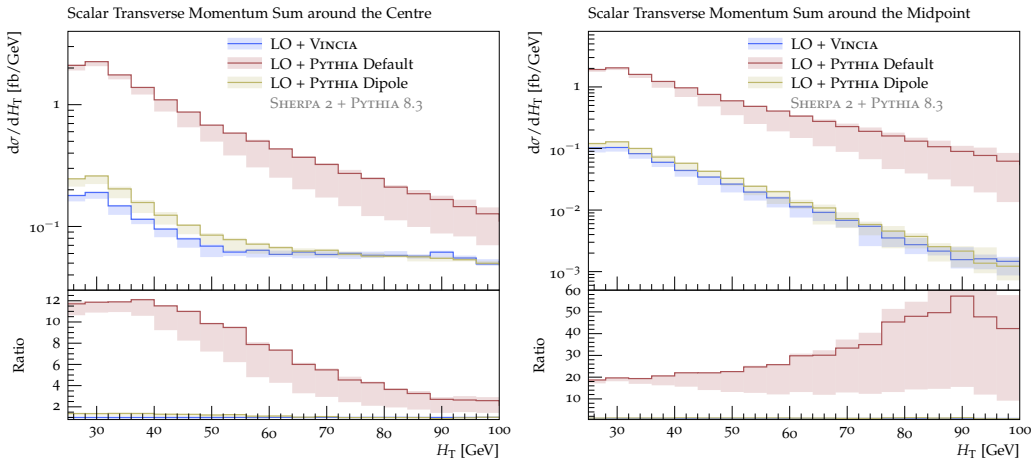


Figure 4: Scalar transverse momentum sum in the central rapidity region (*left*) and around the rapidity midpoint of the tagging jets (*right*) at LO+PS accuracy. The bands are obtained by a variation of the default shower starting scale by a factor of two.

Figure 4 shows the H_{T} distributions in the previously defined central and midpoint regions. As for the third-jet pseudorapidity and transverse-momentum distributions, there is only a minor disagreement between Pythia dipole-recoil shower and Vincia, while Pythia's DGLAP shower generates significantly more radiation in both regions.

For all observables considered here, we also note that the variation of the shower starting scale has a much more pronounced effect on the Pythia default shower than on Vincia or on Pythia when the dipole-recoil option is enabled. Moreover, the starting-scale variation affects the p_{T} distribution of the third jet more than it does the pseudorapidity distribution. This indicates that, while a tailored shower starting scale for the default shower might be able to mimic the phase space-suppression of the dipole/antenna showers to some extent, this would not by itself be sufficient to represent the dipole-antenna emission pattern of the third jet.

We close this subsection by comparing showers off our externally generated Born-level VBF events (i.e., ones generated by Sherpa and passed to Pythia for showering) to showers off internally generated ones (i.e., ones generated by Pythia's `HiggsSM:ff2Hff(t:WW)` and `HiggsSM:ff2Hff(t:ZZ)` processes). This is intended as a cross check for any effects caused by differences in how Pythia treats external vs internal events. For instance, for external events, the external generator is responsible not only for computing the hard cross section but also for setting the shower starting scale, via the HDF5 `scales` dataset (equivalent to the Les Houches `SCALUP` parameter [68, 69]). For our VBF events, the choice made in Sherpa is identical to the factorisation scale eq. (3),

$$\text{Sherpa VBF events:} \qquad \mu_{\mathrm{PS}}^2 \equiv \mu_{\mathrm{F}}^2 = \frac{\hat{H}_{\mathrm{T}}^2}{4} = \frac{1}{4}\left(\sum_j p_{\mathrm{T},j} + \sqrt{M_{\mathrm{H}}^2 + p_{\mathrm{T,H}}^2}\right)^2 .$$

For internally generated VBF events, Pythia's choice of the factorisation scale, and thereby also the shower starting scale, is designed to reflect the off-shellness of the two virtual-boson

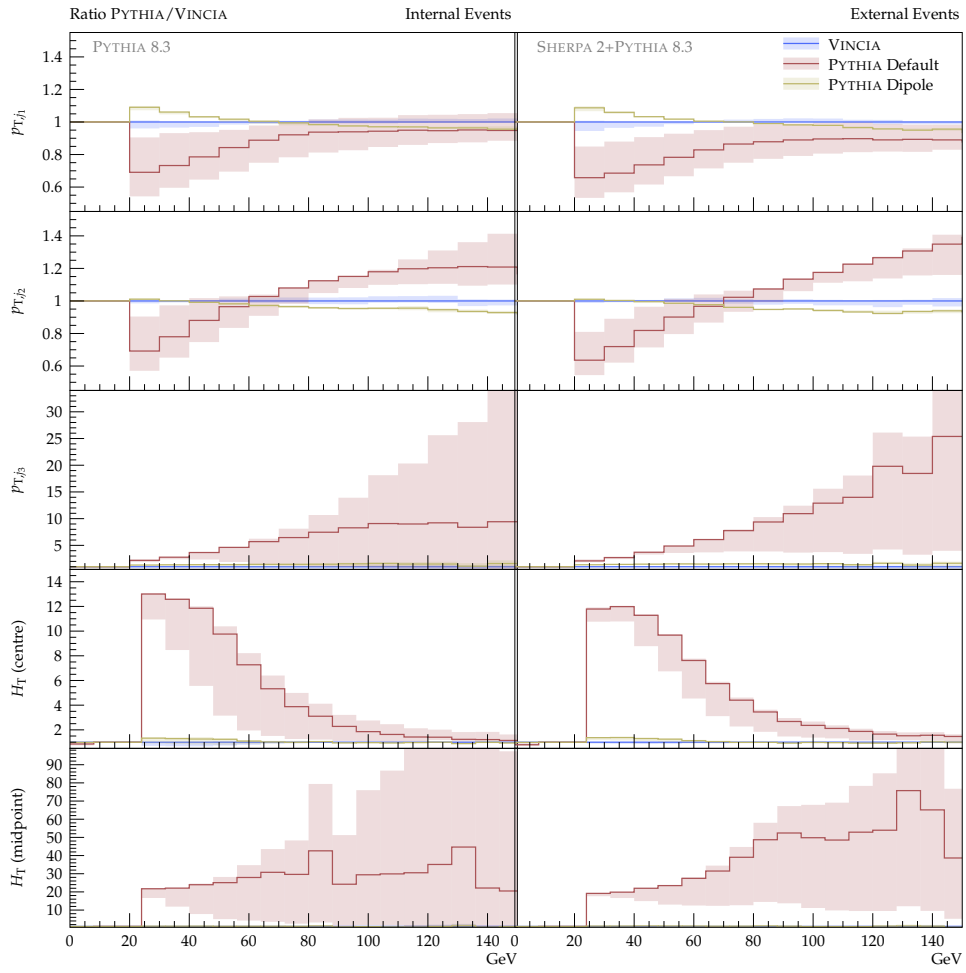


Figure 5: Ratio of Pythia to Vincia at LO+PS accuracy, comparing internal (*left*) and external (*right*) events. The bands are obtained by a variation of the factorisation scale (internal events) and shower starting scale (external events) by a factor of two.

t-channel propagators, cf. eq. (5),

$$\text{Pythia VBF events:} \qquad \mu^2_{\text{PS}} \equiv \mu^2_{\text{F}} = \sqrt{m^2_{\text{T},\text{V}_1} m^2_{\text{T},\text{V}_2}} \equiv \sqrt{(M^2_{\text{V}_1} + p^2_{\text{T},\text{q}_1})(M^2_{\text{V}_2} + p^2_{\text{T},\text{q}_2})}\,.$$

This choice ensures that the factorisation scale and shower starting scale will always be at least of order M^2_{V} even when the outgoing quarks have low $p_{\text{T}} \ll M_{\text{V}}$, while for very large p_{T} values, it asymptotes to the geometric mean of the quark p_{T} values. While the minimum of the Sherpa choice is of the same order, $\mathcal{O}(M_{\text{H}}) \sim \mathcal{O}(M_{\text{V}})$, the large-transverse-momentum limit is considerably larger. The expectation is therefore that, in the absence of matching or merging corrections, Sherpa-generated Born events will lead to higher amounts of hard shower radiation than Pythia-generated ones.

In fig. 5, the ratio of the two Pythia showers to Vincia is shown for the p_{T} and H_{T} spectra using (left) Pythia LO and (right) Sherpa LO events. We immediately note that, in the low-$p_\perp$ limit, the excess of soft radiation generated by Pythia's default shower (red)

persists in both samples. In the high-$p_\perp$ regions, the agreement between the simple shower and the two dipole/antenna options (blue and yellow) tends to be best for PYTHIA's internal hard process. This likely originates from the lower value for the default shower starting scale in PYTHIA, which, as discussed above, imitates the propagator structure of the Born process as closely as possible and hence *should* to some extent set a natural boundary for strongly ordered propagators in the shower. For the dipole/antenna showers, the sensitivity to the starting scale is far milder, as the relevant kinematic information is encoded in the dipole invariant masses independently of the choice of starting scale.

3.2 Next-to-Leading Order Matched

In fig. 6, the POWHEG-matched transverse momentum distributions of the four hardest jets are collected. In comparison to the LO+PS case discussed in the last section, it is directly evident that the Born-jet p_T distributions are in good agreement between all three shower models, including the default PYTHIA one, for which the tagging jet p_T distributions undershoot the VINCIA curve only by an approximately constant factor of order of five per cent. After POWHEG matching, almost perfect agreement is found for the tagging jet transverse momentum distributions obtained with VINCIA and PYTHIA with dipole recoil, as can be seen in fig. 8. The NLO corrections are, however, slightly smaller for the former. The scale choice of the POWHEG emission has only mild effects on all three showers for these tagging-jet observables.

Good agreement is also found between all three shower models for the p_T of the third jet, as shown in the bottom left panel of fig. 6. It must be noted that, again in the case of the PYTHIA default shower, this agreement is subject to appropriately vetoing harder emissions than the POWHEG one, which requires the definition of the POWHEG scale according to the minimal p_T in the event, corresponding to the `POWHEG:pThard = 2` setting, cf. section 2.3.1. Other choices again lead to too hard third jets and heavily increased radiation in the central rapidity region, as can be inferred from the (relative) rapidity distributions of the third jet in the top row of fig. 7, where the importance of a judicious POWHEG scale choice is especially visible. As for the tagging jet spectra, the agreement in both the third-jet transverse momentum and rapidity predictions between VINCIA and the dipole-improved PYTHIA shower is almost perfect, as shown in fig. 9. While the correction (which in this case is essentially a LO matrix-element correction) is positive for VINCIA, it is negative for the dipole-improved PYTHIA shower. Moreover, in the case of VINCIA, this correction affects mostly the high-p_T and the central-rapidity region, whereas for PYTHIA's dipole-improved shower, the correction is negligible at zero rapidity but bigger (and almost) constant at larger rapidities as well as for the transverse momentum.

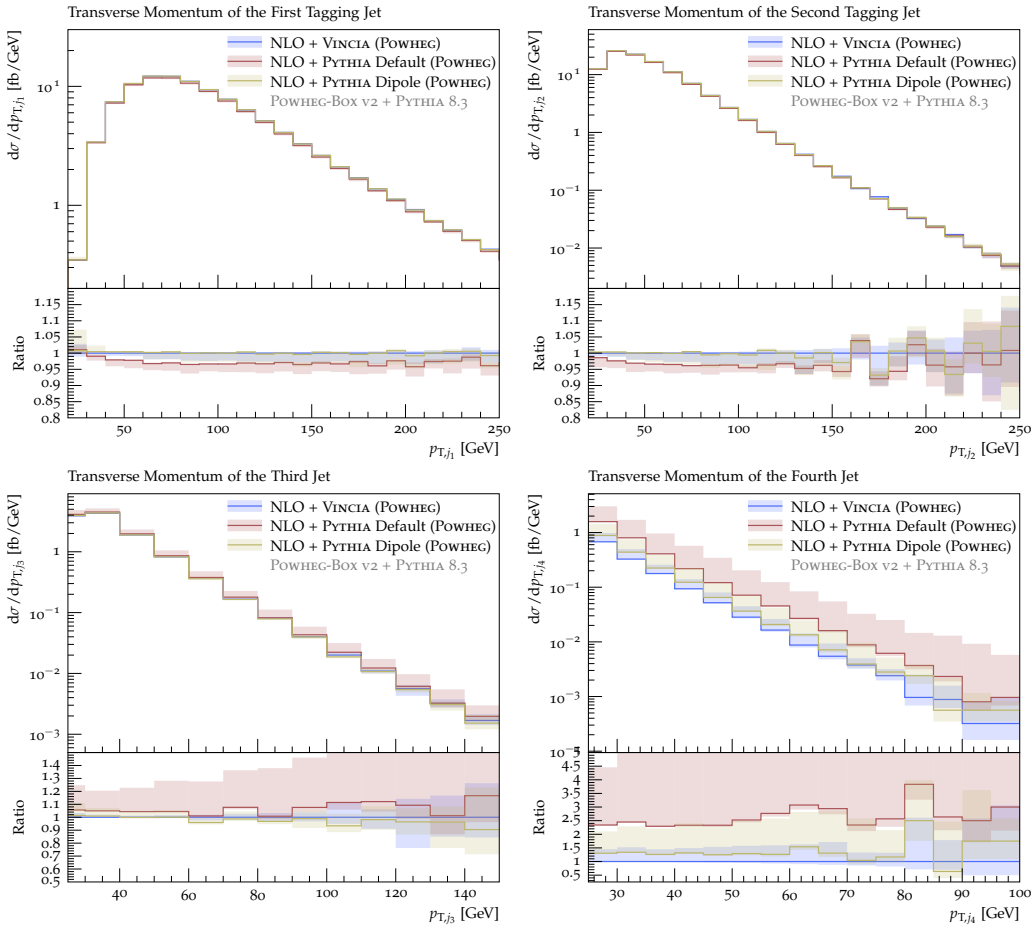


Figure 6: Transverse momentum of the first tagging jet (*top left*), second tagging jet (*top right*), third jet (*top left*), and fourth jet (*top right*) at NLO+PS accuracy in the POWHEG scheme. The bands are obtained by a variation of the hard scale in the vetoed showers as explained in the text.

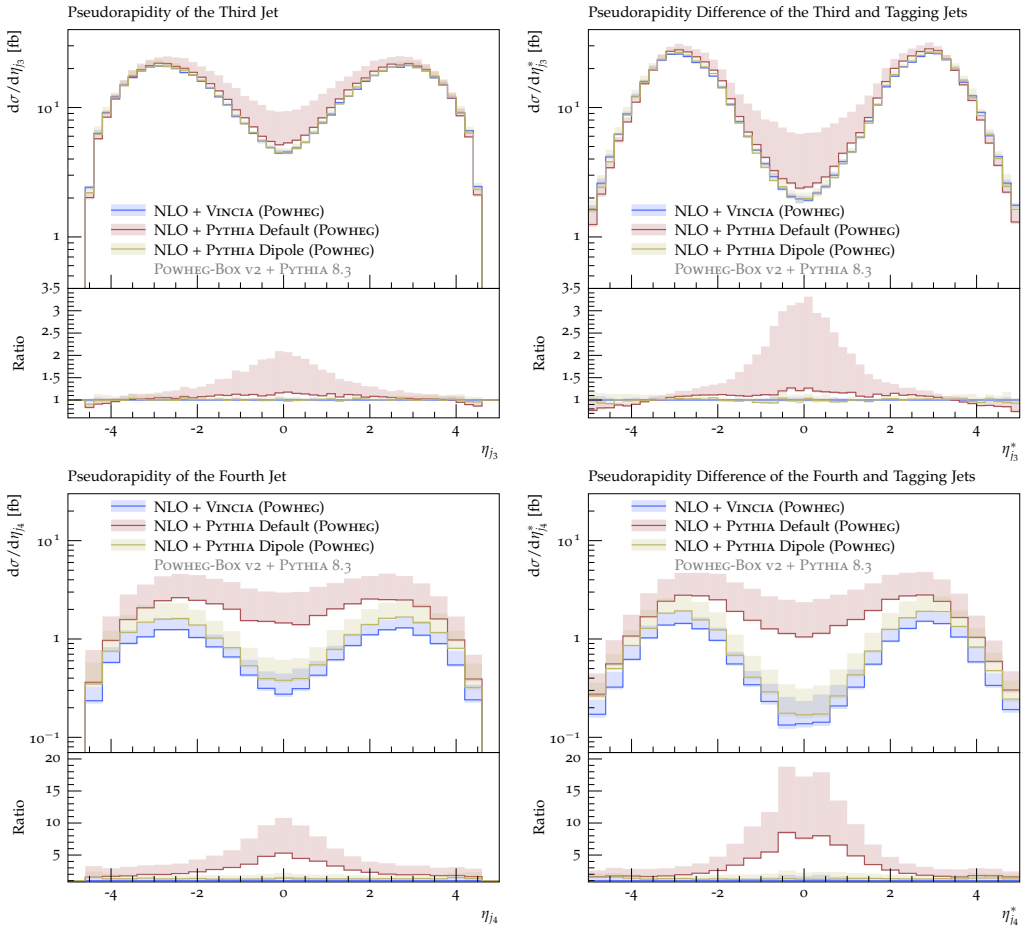


Figure 7: Pseudorapidity (*left column*) and relative rapidity to the tagging jets (*right column*) of the third jet (*top row*) and fourth jet (*bottom row*) at NLO+PS accuracy in the POWHEG scheme. The bands are obtained by a variation of the hard scale in the vetoed showers as explained in the text.

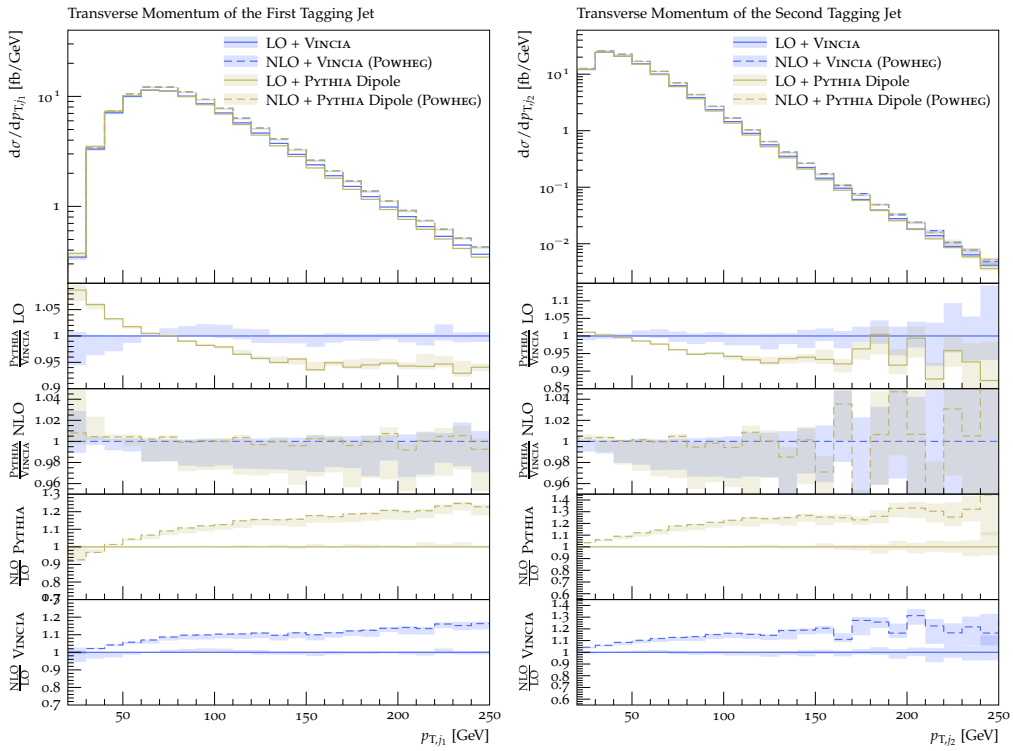


Figure 8: Detailed comparison of the Pythia dipole and Vincia LO+PS and Powheg NLO+PS predictions for the transverse momentum of the first tagging jet (*left*) and the second tagging jet (*right*).

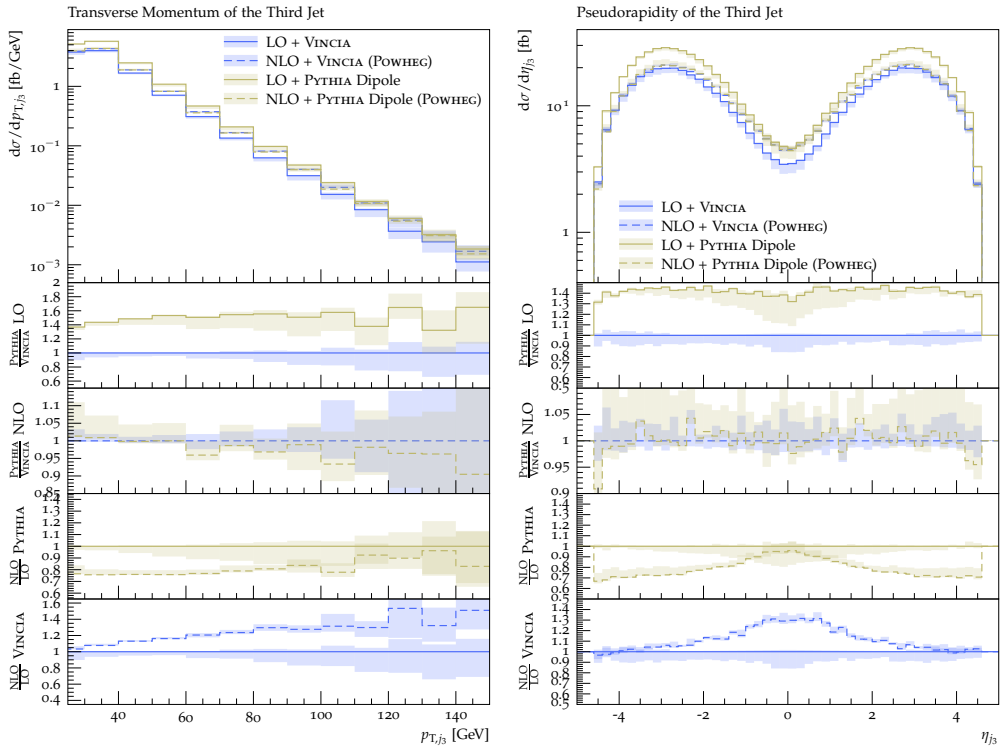


Figure 9: Detailed comparison of the Pythia dipole and Vincia LO+PS and Powheg NLO+PS predictions for the transverse momentum (*left*) and rapidity of the third jet (*right*).

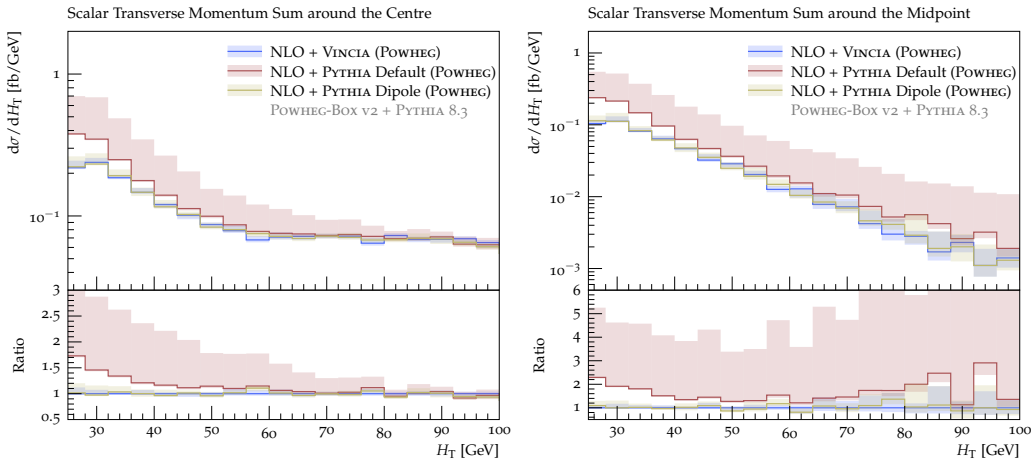


Figure 10: Scalar transverse momentum sum for $|\eta| < 0.5$ (*left*) and around the rapidity midpoint of the tagging jets (*right*) at NLO+PS accuracy in the POWHEG scheme. The bands are obtained by a variation of the hard scale in the vetoed showers as explained in the text.

The bottom right pane in fig. 6 and the bottom row in fig. 7 compare the p_T and (relative) rapidity predictions of the three shower models. While again rather good agreement in these distributions is found for the VINCIA shower and the dipole-improved PYTHIA shower, PYTHIA's default shower produces a harder spectrum, located more in the central rapidity region. Here, it is worthwhile noting that for two-jet POWHEG matching, the emission of the fourth jet is uncorrected in either of the shower models, so that the effects visible in these distributions are solely produced by the showers.

Lastly, fig. 10 shows the scalar transverse momentum for $|\eta| < 0.5$ (left) and around the tagging jet midpoint (right) in the POWHEG NLO+PS scheme. In both distributions, the three shower models produce similar results for $H_T > 40$ GeV, while in the complementary region again only VINCIA and the dipole-improved PYTHIA shower agree. In this soft region, the default PYTHIA shower again predicts more radiation than the other two. As before, a variation of the POWHEG scale choice leads to significant effects in the predictions of PYTHIA's default shower, but has only mild effects on the dipole-improved shower and VINCIA.

3.3 Comparison of Matching and Merging

In figs. 11 to 13, we compare the VINCIA NLO-matched predictions presented in the last section to an $\mathcal{O}(\alpha_S)$ tree-level merged calculation using the CKKW-L scheme implemented for VINCIA. For the latter, we include the exclusive zero-jet and inclusive Sudakov-weighted 1-jet predictions in the plots (dashed lines).

The uncertainty bands of the merged predictions (labelled VINCIA MESS $\mathcal{O}(\alpha_S)$) are obtained by a variation of the shower renormalisation scale as per section 2.2. As VINCIA's merging implementation reweights event samples by a ratio of the strong coupling as used in the shower to the one used in the fixed-order calculation, this variation effectively amounts to an intertwined scale variation of the hard process as well. The uncertainty bands of the NLO-matched calculation are obtained by the variation of the $p_{\perp,\text{hard}}$ scale as in the previous section.

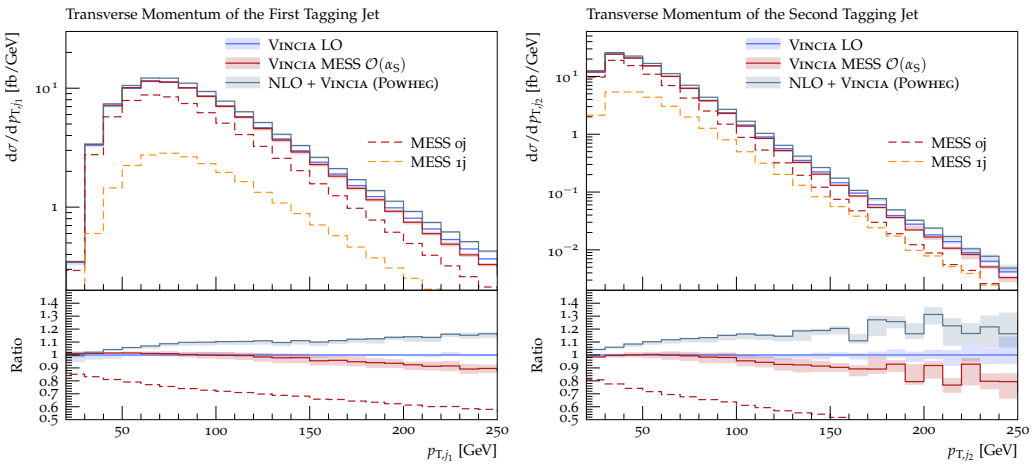


Figure 11: Comparison between LO+PS, POWHEG NLO+PS, and CKKW-L-merged predictions for the transverse momentum of the first (*left*) and second (*right*) tagging jet.

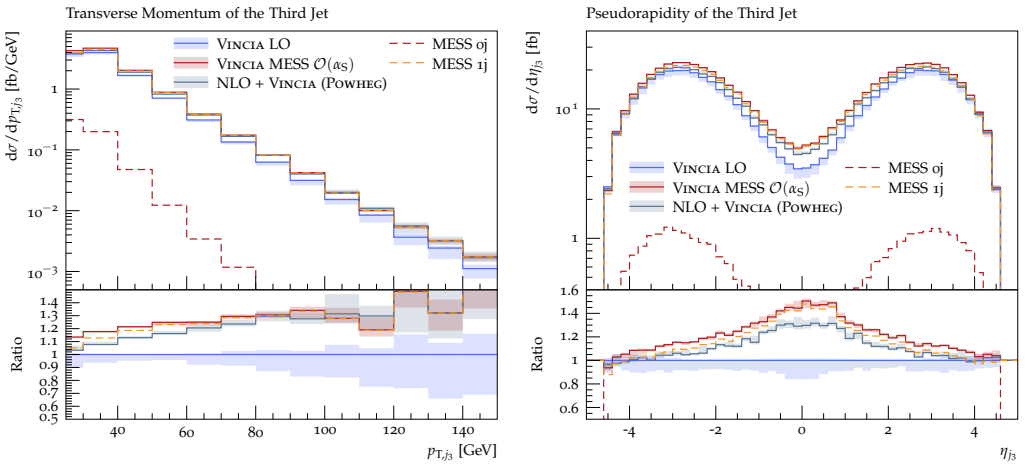


Figure 12: Comparison between LO+PS, POWHEG NLO+PS, and CKKW-L-merged predictions for the transverse momentum (*left*) and pseudorapidity (*right*) of the third jet.

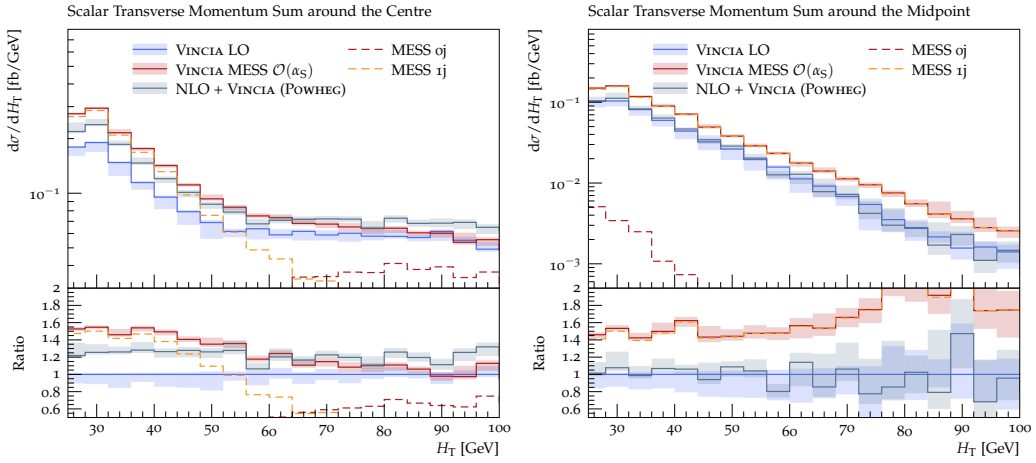


Figure 13: Comparison between LO+PS, POWHEG NLO+PS, and CKKW-L-merged predictions for the scalar transverse momentum sum for $|\eta| < 0.5$ (*left*) and around the pseudorapidity midpoint of the tagging jets (*right*).

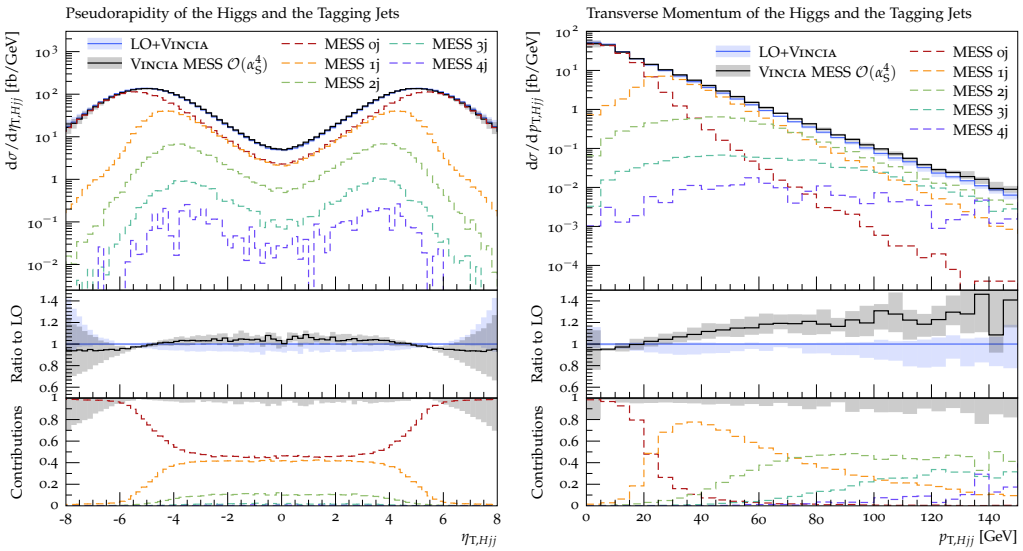


Figure 14: Tree-level merged predictions with up to four additional jets for the pseudorapidity (*left*) and transverse momentum (*right*) of the Higgs and tagging jets system.

Taking into account their respective accuracies, we observe good agreement between the matched and the merged predictions for the transverse momentum and pseudorapidity spectra. We expect the small differences that are visible to trace back mainly to the lack of unitarity in the CKKW-L scheme. This explanation is supported by the fact that the merged calculation overshoots the matched ones and that e.g. for the p_{T,j_3} distribution, the inclusive Sudakov-reweighted 1-jet contribution already agrees in shape and magnitude with the matched distributions, while the exclusive zero-jet contributions only adds to the rate, i.e overall normalisation. In addition, we wish to note again that the mismatch of the POWHEG and VINCIA ordering variables is only treated approximately via the use of vetoed showers, while the correct shower history is taken into account in the merged calculation. Furthermore, we have used two different renormalisation and factorisation scales in the two calculations. Because the renormalisation scale variation in VINCIA's merging affects the renormalisation scale of the hard process, as alluded to above, the renormalisation scale mismatch is covered to some degree by the scale variations in the merging.

The situation is different for the H_T distributions, cf. fig. 13. In the merged calculation, more soft radiation is predicted in the central pseudorapidity region than in the matched one. The distribution is solely governed by the one-jet sample there, while the zero-jet sample contributes significantly above 60 GeV only. In the midpoint region, however, the merged calculation predicts the same shape as the matched one, but with an overall bigger rate. Barely any contribution stems from the exclusive zero-jet sample in this observable. This confirms the properties of the two H_T observables mentioned in section 2.4. When the observable is defined over the central rapidity region, it is sensitive to the radiation of the third jet in the soft region, i.e. for $H_T \lesssim 60$ GeV, but becomes sensitive to the tagging jets in the complementary hard region, i.e. above around 60 GeV. In contrast, defining the observable over the region around the pseudorapidity midpoint of the two tagging jets cleans it from almost all contributions stemming from the Born configuration (only a tiny contribution from soft radiation off the Born survives). Due to this property, the latter of the two definitions is

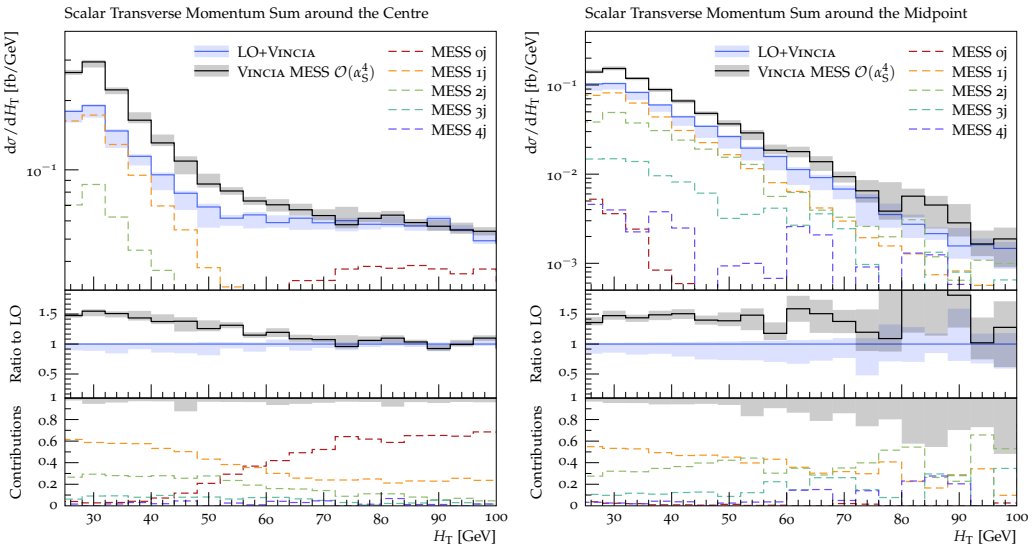


Figure 15: Tree-level merged predictions with up to four additional jets for the scalar transverse momentum sum in the central (*left*) and midpoint (*right*) pseudorapidity region.

particularly suited in the study of the radiation pattern regarding its coherence.

The comparison of NLO matching and $\mathcal{O}(\alpha_S)$ tree-level merging provides a strong cross check of both methods.

3.4 Merged with up to Four Jets

In addition to the one-jet merged calculation of the last section, we here present a tree-level merged calculation with up to four additional jets (i.e., 6 jets in total when counting the tagging jets) using VINCIA's CKKW-L implementation. We consider the effect of additional hard jets on the spectra of the pseudorapidity and transverse momentum of the Higgs plus tagging jets system as well as the herein before mentioned scalar transverse momentum sum in the two pseudorapidity regions. The uncertainty bands of the merged calculation shown in the figures are obtained by a variation of the renormalisation scale prefactors k_R, c.f. section 2.2, in VINCIA's shower and merging, again effectively representing a variation of the renormalisation scale in the hard process as well, cf. section 3.3. As visible from fig. 15, the inclusion of additional hard jets does not change the pseudorapidity spectrum, but increases the rate of the transverse momentum spectrum in the high-p_T region. This correction is exactly what is expected from a multi-jet merged calculation. The dashed lines in fig. 15 represent the different multi-jet contributions to the merged prediction. Again as expected, the Born sample dominates in the low-p_T region and the one-jet sample in the region around 40 GeV, whereas higher multiplicities take over in the harder regions above ~ 70 GeV. It is worth highlighting, however, that, at least in the region 70 GeV $\lesssim p_T \lesssim$ 150 GeV, the two-jet sample dominates with only sub-leading corrections from the three- and four-jet samples.

Figure 14 shows the H_T distributions in the central and midpoint pseudorapidity regions defined in section 2.4. As for the one-jet merged prediction presented in section 3.3, the high-H_T region is dominated by the Born sample, while for small H_T, the samples with additional jets define the shape. Although all samples with additional jets contribute to the central H_T

over the full shown spectrum, the three-jet sample (denoted $1j$ in fig. 14) is the dominant extra-jet sample everywhere. Above approximately 60 GeV, the Born sample becomes the predominant one, highlighting again that this region is sensitive mainly to the tagging jets. Corrections from the multi-jet merging are negligible there.

As before, the situation is different in the midpoint region between the two tagging jets (right-hand pane in fig. 14). There, the Born sample has almost no impact ($< 5\%$) on the H_{T} distribution and the one-jet sample (denoted $1j$ in fig. 14) dominates in the region $\lesssim 70$ GeV, while the two-jet sample (denoted $2j$ in fig. 14) does in the region 70 GeV $\lesssim H_{\mathrm{T}} \lesssim 100$ GeV. This emphasises the finding of the last section that the midpoint H_{T} is clean of contributions from the tagging jets and therefore more relevant in the study of coherence effects in QCD radiation.

3.5 Hadronisation and Multi-Parton Interactions

Although we focused on the parton level throughout this study, we wish to close by estimating the size of non-perturbative corrections arising from hadronisation, fragmentation, and multi-parton interactions. To this end, we employ Pythia's string fragmentation and interleaved MPI model [11] using the default Pythia [52] and Vincia [27] tunes.

Figures 16 to 18 compare Pythia's simple shower and Vincia predictions on the parton level, hadron level, and hadron level with MPIs at LO+PS accuracy. As expected from the cuts employed in our analysis, cf. section 2.4, the inclusion of non-perturbative effects in either of the two simulations has only a negligible effect on most observables studied here, although the discrepancy between the two showers is slightly mitigated. A notable exception are the Vincia predictions for the H_{T} in the two pseudorapidity regions defined in section 2.4, for which the inclusion of MPIs leads to a substantial excess in radiation in the soft region. This means, that in those regions the coherent suppression of radiation by Vincia is overwhelmed by the soft radiation off secondary (non-VBF-like) interactions, at least with our set of cuts. It should be noted here that firstly, this excess is not visible in the distributions obtained with Pythia's simple shower, and secondly, the discrepancy between the simple shower and Vincia overpowers the MPI effect greatly. As such, the inclusion of hadron-level and MPI effects emphasise that Vincia's antenna shower reproduces QCD coherence effects more faithfully than Pythia's simple shower.

4 Conclusion

We have here studied the effect of QCD radiation in VBF Higgs production, focusing in particular on how the coherent emission patterns exhibited by this process are modelled by various parton-shower approaches that are available in the Pythia event generator, and how significant the corrections to that modelling are, from higher fixed-order matrix elements. From a QCD point of view, the main hallmark of VBF is that gluon emission in the central region originates from intrinsically coherent interference between initial- and final-state radiation. In DGLAP-style showers, which are anchored in the collinear limits and treat ISR and FSR separately, this interplay can only be captured at the azimuthally integrated level via angular ordering, while it is a quite natural element in dipole- and antenna-based formalisms, in which initial-final colour flows enter on an equal footing with final-final and initial-initial flows. Hence we would expect the latter (dipole/antenna-style) approaches to offer more ro-

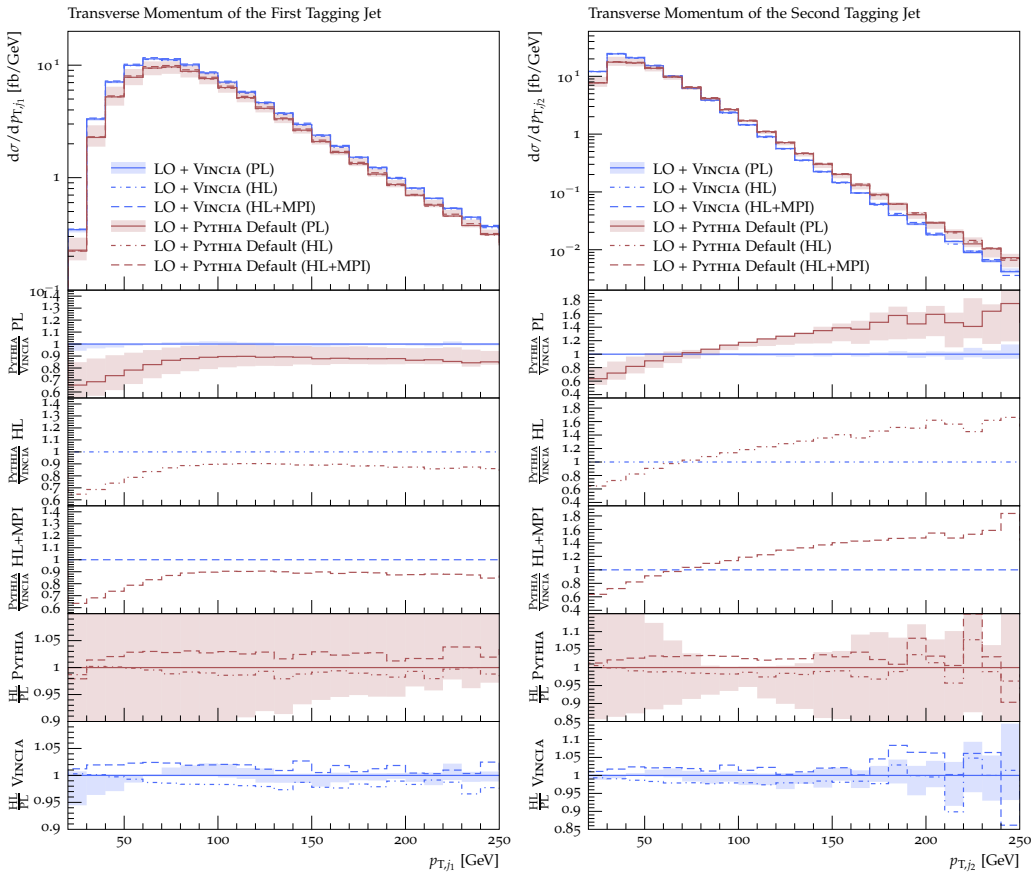


Figure 16: Detailed comparison of Pythia DGLAP and Vincia LO+PS predictions at parton-level, hadron-level, and hadron-level plus MPI for the transverse momentum of the first tagging jet (*left*) and the second tagging jet (*right*).

bust and reliable modelling of the radiation patterns in VBF than the former (DGLAP-based) approaches.

To this end, we have compared the Vincia antenna shower to Pythia's default ("simple") shower, including both its (default) DGLAP and its dipole-improved option ("dipole recoil"). We have shown that at leading order, large discrepancies pertaining to the radiation of additional jets in the central rapidity regions exist between the default Pythia predictions and the ones obtained with the dipole option and Vincia, while the latter two appear more consistent. This effect even concerns observables related to the tagging jets, i.e. those jets which are described by the matrix element and not the shower. We have confirmed that these findings apply to both external (LHA) and internal events.

After matching the showers to the NLO, these discrepancies mostly vanish for observables sensitive to the tagging jets or third jet only, while larger effects remain visible in observables sensitive to higher jet multiplicities. These findings are largely consistent with the ones from an earlier study [8], although it is worth highlighting that the disagreement found for the default Pythia shower is fairly less pronounced here after matching it to the NLO via the Powheg scheme. We consider this to be an effect of a more careful treatment of the ordering-

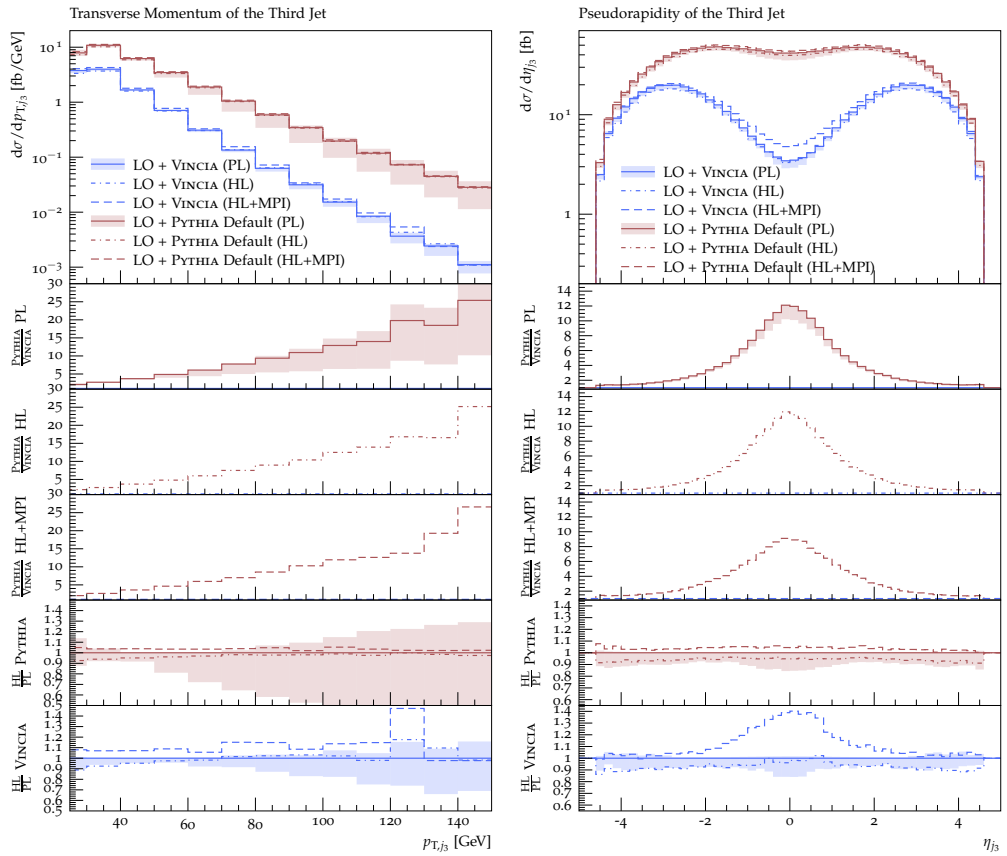


Figure 17: Detailed comparison of Pythia DGLAP and Vincia LO+PS predictions at parton level, hadron level, and hadron-level plus MPI for the transverse momentum (*left*) and pseudorapidity of the third jet (*right*).

variable mismatch between Powheg and Pythia. Based on this, we recommend varying the `POWHEG:pThard` mode contained in the `PowhegHooks` classes to gain an estimate of systematic matching uncertainties. To reduce the uncertainties pertaining to the use of vetoed showers with Powheg samples, a truncated and vetoed shower should be used with both Pythia and Vincia. As alluded to above, such a scheme is not (yet) available for either of the showers considered in the present study.

In addition to NLO matching, we have studied the effect of including higher-multiplicity tree-level matrix elements in the shower via the CKKW-L merging scheme in Vincia. We have confirmed that the NLO-matched and one-jet merged calculations lead to comparable predictions for observables sensitive to the third jet. For a set of inclusive observables, we presented predictions from a tree-level merged calculation at $\mathcal{O}(\alpha_S^4)$. This yields corrections of the order of 20% in the hard tail above around 60 GeV of the transverse momentum spectrum of the Higgs-plus-tagging-jet system. Considering the mild corrections in the ranges studied here, it is evident that the sample with four additional jets (i.e. the $2+4$-jet sample) will contribute significantly only in the very hard tails $H_T \gg 100$ GeV and $p_{\perp,Hjj} \gg 150$ GeV.

Although not the main focus of this study, we have gained a first estimate of non-perturbative corrections on the observables studied here. While we generally found only

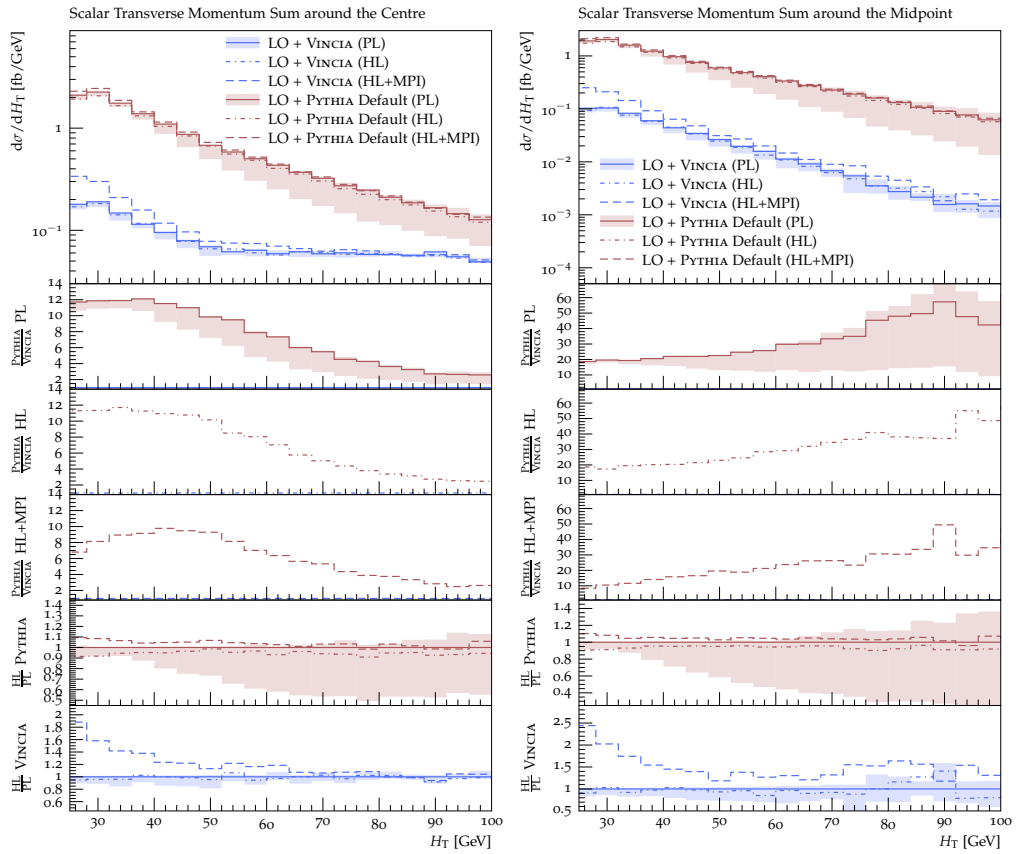


Figure 18: Detailed comparison of Pythia DGLAP and Vincia LO+PS predictions at parton level predictions for the central H_{T} (*left*) and midpoint H_{T} (*right*).

minor changes from the inclusion of hadron-level corrections, the inclusion of MPIs had a relatively more significant effect on Vincia's predictions than on the ones obtained with Pythia's default shower. This affected the rate of radiation in soft as well as central pseudorapidity regions, i.e. precisely the regions in which Vincia predicts a strong coherent suppression, so that the MPI contamination becomes relatively more important.

With this study we also proposed two new observables, the scalar transverse momentum sum in the central pseudorapidity region and around the pseudorapidity midpoint between the two tagging jets. We have shown that both of these observables are sensitive to multi-jet radiation, but highlighted that the former becomes dominated by the tagging jets in the hard region $H_{\mathrm{T}} \gtrsim 60$ G*e*V. As an alternative, we demonstrated that the H_{T} sum around the midpoint between the tagging jets is free of this contamination, with the Born sample only giving a negligible contribution. Due to the strong suppression of radiation in this region, both observables do however receive corrections from the modelling of multi-parton interactions, which would be relevant to study further.

While it has been considered a coherent shower before, this has been the first time that the radiation pattern of the Vincia antenna shower was studied with a dedicated focus on its coherence. At the same time, we have here showcased NLO matching and tree-level merging methods with Vincia, which are both publicly available as of the Pythia 8.306 release.

Acknowledgements

We acknowledge support from the Monash eResearch Centre and eSolutions-Research Support Services through the MonARCH HPC Cluster. This work was further partly funded by the Australian Research Council via Discovery Project DP170100708 — "Emergent Phenomena in Quantum Chromodynamics". CTP is supported by the Monash Graduate Scholarship, the Monash International Postgraduate Research Scholarship, and the J.L. William Scholarship. This research was supported by Fermi National Accelerator Laboratory (Fermilab), a U.S. Department of Energy, Office of Science, HEP User Facility. Fermilab is managed by Fermi Research Alliance, LLC (FRA), acting under Contract No. DE-AC02-07CH11359.

A POWHEG+VINCIA Setup

As mentioned in section 2.3.1, a dedicated vetoed-shower `UserHook` for POWHEG+VINCIA was developed as part of this work and is included in the standard PYTHIA distribution from version 8.306 onwards. At the time of submission of this manuscript, it is included in the file `PowhegHooksVincia.h`, in the directory `include/Pythia8Plugins/`, which also contains the standard `PowhegHooks.h` file. (Note that these two files may be merged into one in a future release; if so, simply omit the corresponding step below.)

Assuming you have a main program that is set up to run POWHEG+PYTHIA (such as the example program `main31.cc` included with PYTHIA), the following changes (highlighted in red) will modify it to run POWHEG+VINCIA:

- Include the `PowhegHooksVincia.h` header file:
  `#include "Pythia8Plugins/PowhegHooksVincia.h"`
  (you can leave any existing `#include "Pythia8Plugins/PowhegHooks.h"` statement; the two will not interfere with each other).

- Replace the POWHEG+PYTHIA user hook pointer by a POWHEG+VINCIA one:
```
shared_ptr<PowhegHooks> powhegHooks;
powhegHooks = make_shared<PowhegHooksVincia>();
pythia.setUserHooksPtr((UserHooksPtr)powhegHooks);
```

In addition, the following settings should be used:

- Switch on VINCIA's showers and allow them to fill all of phase space:
```
PartonShowers:model = 2      # Use Vincia's shower model.
Vincia:pTmaxMatch   = 2      # Power showers (to be vetoed by hook).
```

- Enable shower vetoes via the `PowhegHooksVincia` (same as for `PowhegHooks`):
```
POWHEG:veto         = 1      # Turn shower vetoes on.
```

- Turn QED/EW showers and interleaved resonance decays off:
```
Vincia:ewMode = 0            # Switch off QED/EW showers.
Vincia:interleaveResDec= off # No interleaved resonance decays.
```
  While enabling QED showers (`Vincia:ewMode = 1 | 2`) should not pose any problems in the matching, it is not validated (yet). We recommend against using the EW shower (`Vincia:ewMode = 3`) with the POWHEG matching.

- Since POWHEG-BOX event samples come unpolarised, VINCIA's helicity shower should be turned off (the helicity shower needs a polarised Born state):
```
Vincia:helicityShower  = off # Use helicity-averaged antennae.
```
  We note that VINCIA offers the possibility to polarise Born configurations using matrix elements provided via interfaces to external generators. We have not studied this in the present work.

- In the POWHEG-specific settings, the number of outgoing particles in the Born process is defined as usual, e.g. `=2` for the $2 \to 2$ example in `main31.cc`, or `=3` for the $2 \to 3$ VBF-type processes studied in this work:
```
POWHEG:nFinal  = 3 # Number of outgoing particles in the Born process.
```

- We highly recommend varying the `POWHEG:pThard` mode, for both PYTHIA and VINCIA, to estimate matching systematics. This is how the shaded bands in most of the plots shown in this paper were obtained.
  `POWHEG:pThard  = 2 # Vary (=0,=1,=2) to estimate matching systematics.`
- We also recommend checking all accepted emissions rather than only the first few:
  `POWHEG:vetoCount = 10000`
- The following settings are simply left at their recommended values (the same as for `main31.cmnd`); see the onlin manual section on POWHEG for details:
  ```
  POWHEG:pTemt    = 0
  POWHEG:emitted = 0
  POWHEG:pTdef    = 1
  ```
- For completeness, (we note that we have anyway turned both MPI and QED showers off in this study):
  ```
  POWHEG:MPIveto = 0
  POWHEG:QEDveto = 2
  ```

The event files generated by POWHEG should be provided in exactly the same way as for PYTHIA+POWHEG. If the POWHEG events were generated in several separate batches, for instance, the resulting files can be read as usual, using PYTHIA's "subruns" functionality:

```
! Powheg Subruns.
Beams:frameType      = 4
Main:numberOfSubruns = 3
!-------------------------------------------------------------------
Main:subrun       = 0
Beams:LHEF        = POWHEG-BOX-V2/VBF_H/run/pwgevents-0001.lhe
!-------------------------------------------------------------------
Main:subrun       = 1
Main:LHEFskipInit = on
Beams:LHEF        = POWHEG-BOX-V2/VBF_H/run/pwgevents-0002.lhe
!-------------------------------------------------------------------
Main:subrun       = 2
Main:LHEFskipInit = on
Beams:LHEF        = POWHEG-BOX-V2/VBF_H/run/pwgevents-0003.lhe
```

B VINCIA CKKW-L Setup

Since PYTHIA version 8.304, the release is shipped with VINCIA's own implementation of the CKKW-L merging technique, suitably modified for sector showers.

In the spirit of the last section, let us again assume you have a main program running CKKW-L merging with PYTHIA's default ("simple") shower. (We note that this is a hypothetical setup for the purpose of this study, as the default merging implementation in PYTHIA 8.3 does not handle VBF processes. An algorithmic fix is planned for PYTHIA version 8.307 or later.) The following changes are needed to alter it to run VINCIA's CKKW-L merging instead, with changes again highlighted in red.

- Turn VINCIA and its sector showers on[6]:
```
PartonShowers:model  = 2     # Use Vincia's shower model.
Vincia:sectorShowers = on    # Turn sector showers on.
```

- Disable VINCIA components that are not (yet) handled by the merging:
```
Vincia:ewMode = 0            # Switch off QED/EW showers.
Vincia:interleaveResDec= off # No interleaved resonance decays.
Vincia:helicityShower  = off # Use helicity-averaged antennae.
```
These three limitations are intended to be temporary and may be lifted in future updates; users are encouraged to check for changes mentioning VINCIA's merging implementation in the Update History section of PYTHIA's HTML manual in releases from 8.307 onwards.

- Enable the merging machinery and set the merging scale definition (in this study, all event samples were regulated by a k_{T} cut, so k_{T}-merging is turned on):
```
Merging:doMerging   = on     # Turn merging machinery on.
Merging:doKTMerging = on     # Set kT as merging scale.
```

- Set the merging scale to the desired value in GeV (note that the cuts on the event samples should be more inclusive than the ones in the merging!):
```
Merging:TMS     = 20         # Value of the merging scale in GeV.
```

- Replace the `Process` string by one obeying VINCIA's syntax, i.e. encased in curly brackets and with whitespaces between particles, and switch the dedicated VBF treatment on:
```
Merging:process = { p p > h0 j j } # Define the hard process.
Vincia:mergeVBF = on               # Enable merging in VBF systems.
```

- Set the number of additional jets with respect to the Born process (e.g. for the VBF process considered here, the number of *additional* jets is 4, while the *total* number of jets is 6):
```
Merging:nJetMax = 4    # Merge samples with up to 4 additional jets.
```

References

[1] D. de Florian *et al.*, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector* **2/2017** (2016), doi:10.23731/CYRM-2017-002, 1610.07922.

[2] F. A. Dreyer and A. Karlberg, *Vector-Boson Fusion Higgs Production at Three Loops in QCD*, Phys. Rev. Lett. **117**(7), 072001 (2016), doi:10.1103/PhysRevLett.117.072001, 1606.00840.

[3] R. V. Harlander, J. Vollinga and M. M. Weber, *Gluon-Induced Weak Boson Fusion*, Phys. Rev. D **77**, 053010 (2008), doi:10.1103/PhysRevD.77.053010, 0801.3355.

[4] M. Cacciari, F. A. Dreyer, A. Karlberg, G. P. Salam and G. Zanderighi, *Fully Differential Vector-Boson-Fusion Higgs Production at Next-to-Next-to-Leading Order*, Phys.

[6] We note that as of now, sector showers are on per default in VINCIA and this flag is listed here only for completeness.

Rev. Lett. **115**(8), 082002 (2015), doi:10.1103/PhysRevLett.115.082002, [Erratum: Phys.Rev.Lett. 120, 139901 (2018)], 1506.02660.

[5] J. Cruz-Martinez, T. Gehrmann, E. W. N. Glover and A. Huss, *Second-order QCD effects in Higgs boson production through vector boson fusion*, Phys. Lett. B **781**, 672 (2018), doi:10.1016/j.physletb.2018.04.046, 1802.02445.

[6] T. Liu, K. Melnikov and A. A. Penin, *Nonfactorizable QCD Effects in Higgs Boson Production via Vector Boson Fusion*, Phys. Rev. Lett. **123**(12), 122002 (2019), doi:10.1103/PhysRevLett.123.122002, 1906.10899.

[7] M. Ciccolini, A. Denner and S. Dittmaier, *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC*, Phys. Rev. D **77**, 013002 (2008), doi:10.1103/PhysRevD.77.013002, 0710.4749.

[8] B. Jäger, A. Karlberg, S. Plätzer, J. Scheller and M. Zaro, *Parton-shower effects in Higgs production via Vector-Boson Fusion*, Eur. Phys. J. C **80**(8), 756 (2020), doi:10.1140/epjc/s10052-020-8326-7, 2003.12435.

[9] A. Buckley *et al.*, *A comparative study of Higgs boson production from vector-boson fusion* (2021), 2105.11399.

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

[11] T. Sjöstrand and P. Z. Skands, *Transverse-momentum-ordered showers and interleaved multiple interactions*, Eur. Phys. J. C **39**, 129 (2005), doi:10.1140/epjc/s2004-02084-y, hep-ph/0408302.

[12] R. Corke and T. Sjöstrand, *Interleaved Parton Showers and Tuning Prospects*, JHEP **03**, 032 (2011), doi:10.1007/JHEP03(2011)032, 1011.1759.

[13] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen and P. Z. Skands, *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191**, 159 (2015), doi:10.1016/j.cpc.2015.01.024, 1410.3012.

[14] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07**, 079 (2014), doi:10.1007/JHEP07(2014)079, 1405.0301.

[15] 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**, 043 (2010), doi:10.1007/JHEP06(2010)043, 1002.2581.

[16] B. Cabouat and T. Sjöstrand, *Some Dipole Shower Studies*, Eur. Phys. J. C **78**(3), 226 (2018), doi:10.1140/epjc/s10052-018-5645-z, 1710.00391.

[17] J. Bellm *et al.*, *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76**(4), 196 (2016), doi:10.1140/epjc/s10052-016-4018-8, 1512.01178.

[18] M. Rauch and S. Plätzer, *Parton Shower Matching Systematics in Vector-Boson-Fusion WW Production*, Eur. Phys. J. C **77**(5), 293 (2017), doi:10.1140/epjc/s10052-017-4860-3, 1605.07851.

[19] B. Jäger, F. Schissler and D. Zeppenfeld, *Parton-shower effects on Higgs boson production via vector-boson fusion in association with three jets*, JHEP **07**, 125 (2014), doi:10.1007/JHEP07(2014)125, 1405.6950.

[20] R. Covarelli, M. Pellen and M. Zaro, *Vector-Boson scattering at the LHC: Unraveling the electroweak sector*, Int. J. Mod. Phys. A **36**(16), 2130009 (2021), doi:10.1142/S0217751X2130009X, 2102.10991.

[21] D. Buarque *et al.*, *Vector Boson Scattering Processes: Status and Prospects* (2021), 2106.01393.

[22] *Modelling of the vector boson scattering process $pp \to W^{\pm}W^{\pm}jj$ in Monte Carlo generators in ATLAS* (2019).

[23] G. Aad *et al.*, *Measurements of Higgs Bosons Decaying to Bottom Quarks from Vector Boson Fusion Production with the ATLAS Experiment at $\sqrt{s}$ = 13 TeV* (2020), 2011.08280.

[24] G. Aad *et al.*, *Search for Higgs boson production in association with a high-energy photon via vector-boson fusion with decay into bottom quark pairs at $\sqrt{s}$=13 TeV with the ATLAS detector* (2020), 2010.13651.

[25] A. M. Sirunyan *et al.*, *Search for invisible decays of a Higgs boson produced through vector boson fusion in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, Phys. Lett. B **793**, 520 (2019), doi:10.1016/j.physletb.2019.04.025, 1809.05937.

[26] A. M. Sirunyan *et al.*, *Measurements of Higgs boson production cross sections and couplings in the diphoton decay channel at $\sqrt{s}$ = 13 TeV* (2021), 2103.06956.

[27] H. Brooks, C. T. Preuss and P. Skands, *Sector Showers for Hadron Collisions*, JHEP **07**, 032 (2020), doi:10.1007/JHEP07(2020)032, 2003.00702.

[28] G. Gustafson and U. Pettersson, *Dipole Formulation of QCD Cascades*, Nucl. Phys. B **306**, 746 (1988), doi:10.1016/0550-3213(88)90441-5.

[29] B. Andersson, G. Gustafson and C. Sjogren, *Comparison of the dipole cascade model versus O(alpha-s**2) matrix elements and color interference in e+ e- annihilation*, Nucl. Phys. B **380**, 391 (1992), doi:10.1016/0550-3213(92)90250-F.

[30] G. Gustafson, *Multiplicity distributions in QCD cascades*, Nucl. Phys. B **392**, 251 (1993), doi:10.1016/0550-3213(93)90203-2.

[31] C. Friberg, G. Gustafson and J. Hakkinen, *Color connections in e+ e- annihilation*, Nucl. Phys. B **490**, 289 (1997), doi:10.1016/S0550-3213(97)00064-3, hep-ph/9604347.

[32] S. Platzer and S. Gieseke, *Coherent Parton Showers with Local Recoils*, JHEP **01**, 024 (2011), doi:10.1007/JHEP01(2011)024, 0909.5593.

[33] M. Dasgupta, F. A. Dreyer, K. Hamilton, P. F. Monni and G. P. Salam, *Logarithmic accuracy of parton showers: a fixed-order study*, JHEP **09**, 033 (2018), doi:10.1007/JHEP09(2018)033, [Erratum: JHEP 03, 083 (2020)], 1805.09327.

[34] J. R. Forshaw, J. Holguin and S. Plätzer, *Building a consistent parton shower*, JHEP **09**, 014 (2020), doi:10.1007/JHEP09(2020)014, 2003.06400.

[35] J. Holguin, J. R. Forshaw and S. Plätzer, *Improvements on dipole shower colour*, Eur. Phys. J. C **81**(4), 364 (2021), doi:10.1140/epjc/s10052-021-09145-1, 2011.15087.

[36] M. Dasgupta, F. A. Dreyer, K. Hamilton, P. F. Monni, G. P. Salam and G. Soyez, *Parton showers beyond leading logarithmic accuracy*, Phys. Rev. Lett. **125**(5), 052002 (2020), doi:10.1103/PhysRevLett.125.052002, 2002.11114.

[37] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *QCD matrix elements + parton showers*, JHEP **11**, 063 (2001), doi:10.1088/1126-6708/2001/11/063, hep-ph/0109231.

[38] L. Lönnblad, *Correcting the color dipole cascade model with fixed order matrix elements*, JHEP **05**, 046 (2002), doi:10.1088/1126-6708/2002/05/046, hep-ph/0112284.

[39] L. Lönnblad and S. Prestel, *Matching Tree-Level Matrix Elements with Interleaved Showers*, JHEP **03**, 019 (2012), doi:10.1007/JHEP03(2012)019, 1109.4829.

[40] H. Brooks and C. T. Preuss, *Efficient multi-jet merging with the Vincia sector shower*, Comput. Phys. Commun. **264**, 107985 (2021), doi:10.1016/j.cpc.2021.107985, 2008.09468.

[41] E. Bothmann *et al.*, *Event Generation with Sherpa 2.2*, SciPost Phys. **7**(3), 034 (2019), doi:10.21468/SciPostPhys.7.3.034, 1905.09127.

[42] S. Höche, S. Prestel and H. Schulz, *Simulation of Vector Boson Plus Many Jet Final States at the High Luminosity LHC*, Phys. Rev. D **100**(1), 014024 (2019), doi:10.1103/PhysRevD.100.014024, 1905.05120.

[43] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump and C. P. Yuan, *New parton distribution functions from a global analysis of quantum chromodynamics*, Phys. Rev. D **93**(3), 033006 (2016), doi:10.1103/PhysRevD.93.033006, 1506.07443.

[44] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr and G. Watt, *LHAPDF6: parton density access in the LHC precision era*, Eur. Phys. J. C **75**, 132 (2015), doi:10.1140/epjc/s10052-015-3318-8, 1412.7420.

[45] R. D. Ball *et al.*, *Parton distributions with LHC data*, Nucl. Phys. B **867**, 244 (2013), doi:10.1016/j.nuclphysb.2012.10.003, 1207.1303.

[46] R. D. Ball, V. Bertone, S. Carrazza, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland and J. Rojo, *Parton distributions with QED corrections*, Nucl. Phys. B **877**, 290 (2013), doi:10.1016/j.nuclphysb.2013.10.010, 1308.0598.

[47] T. Gleisberg and S. Hoeche, *Comix, a new matrix element generator*, JHEP **12**, 039 (2008), doi:10.1088/1126-6708/2008/12/039, 0808.3674.

[48] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11**, 040 (2004), doi:10.1088/1126-6708/2004/11/040, hep-ph/0409146.

[49] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11**, 070 (2007), doi:10.1088/1126-6708/2007/11/070, 0709.2092.

[50] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP **02**, 037 (2010), doi:10.1007/JHEP02(2010)037, 0911.5299.

[51] P. Skands, B. Webber and J. Winter, *QCD Coherence and the Top Quark Asymmetry*, JHEP **07**, 151 (2012), doi:10.1007/JHEP07(2012)151, 1205.1466.

[52] P. Skands, S. Carrazza and J. Rojo, *Tuning PYTHIA 8.1: the Monash 2013 Tune*, Eur. Phys. J. C **74**(8), 3024 (2014), doi:10.1140/epjc/s10052-014-3024-y, 1404.5630.

[53] S. Amoroso *et al.*, *Les Houches 2019: Physics at TeV Colliders: Standard Model Working Group Report*, In *11th Les Houches Workshop on Physics at TeV Colliders: PhysTeV Les Houches* (2020), 2003.01700.

[54] S. Catani, B. R. Webber and G. Marchesini, *QCD coherent branching and semiinclusive processes at large x*, Nucl. Phys. B **349**, 635 (1991), doi:10.1016/0550-3213(91)90390-J.

[55] S. Hoeche, F. Krauss, M. Schönherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP **09**, 049 (2012), doi:10.1007/JHEP09(2012)049, 1111.1220.

[56] P. Nason and B. Webber, *Next-to-Leading-Order Event Generators*, Ann. Rev. Nucl. Part. Sci. **62**, 187 (2012), doi:10.1146/annurev-nucl-102711-094928, 1202.1251.

[57] R. Corke and T. Sjöstrand, *Improved Parton Showers at Large Transverse Momenta*, Eur. Phys. J. C **69**, 1 (2010), doi:10.1140/epjc/s10052-010-1409-0, 1003.2384.

[58] T. Plehn, D. Rainwater and P. Z. Skands, *Squark and gluino production with jets*, Phys. Lett. B **645**, 217 (2007), doi:10.1016/j.physletb.2006.12.009, hep-ph/0510144.

[59] S. Gieseke, P. Stephens and B. Webber, *New formalism for QCD parton showers*, JHEP **12**, 045 (2003), doi:10.1088/1126-6708/2003/12/045, hep-ph/0310083.

[60] K. Hamilton, P. Richardson and J. Tully, *A Modified CKKW matrix element merging approach to angular-ordered parton showers*, JHEP **11**, 038 (2009), doi:10.1088/1126-6708/2009/11/038, 0905.3072.

[61] S. Hoeche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **05**, 053 (2009), doi:10.1088/1126-6708/2009/05/053, 0903.1219.

[62] S. Hoche, F. Krauss, M. Schonherr and F. Siegert, *NLO matrix elements and truncated showers*, JHEP **08**, 123 (2011), doi:10.1007/JHEP08(2011)123, 1009.1127.

[63] P. Nason and C. Oleari, *Generation cuts and Born suppression in POWHEG* (2013), 1303.3922.

[64] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04**, 063 (2008), doi:10.1088/1126-6708/2008/04/063, 0802.1189.

[65] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, Eur. Phys. J. C **72**, 1896 (2012), doi:10.1140/epjc/s10052-012-1896-2, 1111.6097.

[66] A. Buckley, J. Butterworth, D. Grellscheid, H. Hoeth, L. Lönnblad, J. Monk, H. Schulz and F. Siegert, *Rivet user manual*, Comput. Phys. Commun. **184**, 2803 (2013), doi:10.1016/j.cpc.2013.05.021, 1003.0694.

[67] C. Bierlich *et al.*, *Robust Independent Validation of Experiment and Theory: Rivet version 3*, SciPost Phys. **8**, 026 (2020), doi:10.21468/SciPostPhys.8.2.026, 1912.05451.

[68] E. Boos *et al.*, *Generic User Process Interface for Event Generators*, In *2nd Les Houches Workshop on Physics at TeV Colliders* (2001), hep-ph/0109068.

[69] J. Alwall *et al.*, *A Standard format for Les Houches event files*, Comput. Phys. Commun. **176**, 300 (2007), doi:10.1016/j.cpc.2006.11.010, hep-ph/0609017.

9

Conclusions

In this thesis, work on Monte Carlo event generators with a special focus on parton showers was presented. Monte Carlo event generators are indispensable tools for collider phenomenology, *i.e.*, the study of particle interactions in high-energy collisions, as they facilitate tests of our understanding of the Standard Model as well as searches for previously unexplained new-physics phenomena. This thesis addressed two pressing issues in state-of-the-art event generators: the increasing computational overhead related to event generation and the need for higher-precision simulations.

In chapter 2, a brief introduction to Quantum Field Theory and the Standard Model, the theoretical framework underpinning modern particle physics, was given. After reviewing the weak and strong interactions as local gauge theories, the master formula for the calculation of cross sections in collider experiments was highlighted and their importance in particle physics phenomenology detailed. Subsequently, chapter 3 covered the methods employed in Monte Carlo event generators and contained a detailed overview over existing implementations. Special attention was paid to an in-depth description of parton showers as well as matching and merging techniques. The discussion in chapters 2 and 3, in particular the latter, provides the foundations of the remaining chapters of this thesis, in which, based on published works, the core results are presented.

The full-fledged implementation of so-called sector showers in the VINCIA antenna shower has been presented in chapter 4, extending a previous proof-of-concept implementation of final-state sector showers [327] to initial-state radiation and showers in coloured resonance decays. The most prominent feature of sector showers is that they are maximally bijective, meaning that the minimal possible number of so-called shower histories is produced during the shower evolution. In best cases (when only gluons are emitted), any given multi-parton configuration can be traced back through a single sequence of intermediate states. For configurations with multiple quark pairs, the number of histories is given by the number of viable quark pair permutations. For each permutation, the history is again unique. The sector-shower implementation in VINCIA has been made public with the PYTHIA 8.304 release.

A CKKW-L-based tree-level multi-jet merging scheme tailored to VINCIA's sector showers was introduced in chapter 5. The maximally bijective nature of sector showers reduces the factorial scaling of the number of shower histories to an effective linear scaling with the number of final-state particles. As the number of histories drives the efficiency and memory footprint of CKKW-L-style merging schemes, the use of sector showers drastically improves the memory allocation and event-generation time to an approximate

constant scaling with the particle multiplicity. This was demonstrated in vector boson production at the LHC with up to nine additional jets. The new merging scheme, dubbed MESS as an abbreviation for matrix elements plus sector showers, has been implemented in VINCIA and was made publicly available with the PYTHIA 8.304 release.

Another major bottleneck in state-of-the-art event generation is the evaluation of one-loop matrix elements, which is typically done by interfacing automated one-loop providers. In chapter 6, a new C++ interface to the MCFM parton-level event generator has been presented. It gives access to MCFM's extensive library of analytic one-loop matrix elements in a way that they can be used in any event generation framework. The interface was tested both with the PYTHIA and SHERPA event generators. For a large number of processes $pp \to n$ with $n \leq 4$, it was shown that the corresponding one-loop matrix elements are evaluated at least a factor ten faster using the interface than with automated tools. In a typical next-to-leading order merged setup of SHERPA, it was demonstrated that the use of the new interface decreases the particle-level event-generation time by 30%–80%. The interface has been made publicly available in a pre-release of MCFM 10.0 and will be part of the next major MCFM release. It is the first publicly-available analytic one-loop provider.

The precision frontier has been addressed in chapter 7, where a novel method to match parton showers to next-to-next-to-leading order calculations has been outlined. Facilitated by the use of VINCIA's sector showers in conjunction with iterated (next-to-)leading-order matrix-element corrections and the MCFM interface, the method allows, for the first time, the fully-differential matching at this precision similar in style to the POWHEG method. As a proof of feasibility, all components of the matching scheme have been implemented and validated in VINCIA for the simplest case of $e^+e^- \to 2j$.

In chapter 8, VINCIA's sector shower was employed in a study of Higgs boson production via vector boson fusion. The distinct colour topology of these processes poses a non-trivial environment for parton showers, as the radiative suppression in the central rapidity region has to be modelled appropriately. It has been demonstrated that both VINCIA's sector shower and PYTHIA's dipole-improved shower reproduce this coherent suppression more faithfully than PYTHIA's default transverse-momentum-ordered DGLAP shower. The reliability of these predictions was tested as higher-order perturbative corrections are incorporated via the POWHEG next-to-leading order matching scheme in PYTHIA and VINCIA and via multi-jet merging with up to four additional jets using VINCIA's CKKW-L implementation. The latter was suitably extended to deal with processes with jets on the Born level. With a judicious and careful setup, all three showers can be brought to agreement for observables sensitive to radiation up to the matched order. Systematic uncertainties pertaining to the use of the POWHEG matching scheme with the PYTHIA framework have been addressed. This study provided a strong proof that VINCIA can be employed for the simulation of non-trivial processes in real-life setups. In addition, for the first time, an implementation of vetoed showers for POWHEG matching with VINCIA has been provided. It has been made publicly available with the PYTHIA 8.306 release.

With the implementations detailed in this thesis, PYTHIA 8.3 offers three different, dedicated parton-shower models with associated matching and merging algorithms: PYTHIA's default DGLAP shower, the VINCIA antenna shower, and the DIRE dipole shower. The availability of three independent showers offers great potential for cross checks and the assessment of uncertainties.

The ever-increasing demand for detailed simulations of collider processes drives the development of Monte Carlo tools regarding their efficiency as well as their accuracy and precision. With the advent of the high-luminosity LHC, the precision constraints

and efficiency bottlenecks of current event-generation tools will be more pressing than ever before. The unprecedented amount of data not only of the high-luminosity LHC but also future collider experiments will only be of its anticipated value if it can be compared to theoretical predictions with high statistical significance. This issue is well known and efforts to tackle these challenges are ongoing. In this thesis, a small subset of these efforts was presented, focussed on the improvement of parton showers and their systematic combination with fixed-order calculations.

An appreciably important facet of parton-shower algorithms today and in the future will be the formal assessment of their logarithmic accuracy. Although this is a highly active field with rapid developments, no attempt was made within the scope of this thesis to determine the formal accuracy of Vincia's sector showers. This cannot be neglected in future developments. While sector showers may complicate the analytic assessment of their accuracy through the use of non-trivial step functions, they offer the possibility for simple algorithms to address shortcomings. It is, for instance, known that dipole/antenna showers assign the wrong subleading-colour factor to regions in which emissions have commensurate transverse momentum but disparate angles [288]. To this end, for instance the assignment of colour factors along the lines proposed in [281, 320, 321] is straight-forward as a tentative post-branching state is anyway constructed for every trial branching, from which the angular ordering of the emission can be read off.

With the availability of full-fledged sector showers a series of further developments is conceivable. Currently, sector showers are constructed in such a way that emissions are restricted to their respective phase-space regions via explicit sector vetoes. As this amounts to a potentially large oversampling of the branching phase space, an immediate improvement is therefore the refinement of sector sampling methods. This will directly translate into a speed-up of up to a factor of three, as this is the number by which the phase space is on average oversampled. Work in this direction is currently ongoing.

In future work, the sector merging approach may be extended towards next-to-leading order merging. Given the immense efficiency gains observed when switching from conventional CKKW-L tree-level merging to sectorised CKKW-L merging, it is not unreasonable to expect comparable gains at NLO. As NLO multi-jet merging de-facto represents the state-of-the-art for many observables and processes studied at the LHC, this will be one of the key requirements for shaping Vincia and its sector showers up for the use by experimental collaborations. To this end, not only the implementation of existing schemes, such as UNLOPS or MINLO, but also the generalisation of NLO matrix-element corrections should be considered. The latter will be an important ingredient of a generalised NNLO matching framework in Vincia.

To date, a wide variety of standard-model processes can be calculated fully-differentially at NNLO accuracy. Despite the existence of a few techniques to combine these with parton-shower simulations, NNLO matching has not reached the same maturity as NLO matching, which by now is well-established. Facilitated by the use of the antenna framework in both the fixed-order calculation and the shower, the proof-of-concept NNLO matching presented in chapter 7 can be upgraded to a full-fledged framework, including hadron collisions. This will require the inclusion of second-order antenna functions not only in Vincia's FSR but also ISR evolution, at least for the first (double-)emission. In this context, it will be interesting to compare this NNLO matching scheme to existing, merging-based approaches such as UN2LOPS or MINNLO$_{\text{PS}}$, with a particular focus on matching-scheme uncertainties.

Extending the last point, a complete second-order shower evolution may be considered in future developments. This will constitute an important aspect of endeavours to reach higher logarithmic accuracies with sector showers. As such, it goes hand in hand with

the assessment of their formal accuracy and efforts to correct shortcomings of the existing strongly-ordered components.

In chapter 6, it was noted that the calculation of real corrections constitutes one of the biggest bottlenecks in NLO calculations. To mitigate this, an extension of the MCFM interface to tree-level matrix elements may be considered in the future, as this will lead to comparable speed-ups over automated matrix-element generators as seen for one-loop matrix elements. As NNLO calculations will become more important and more widely employed, the inclusion of MCFM's two-loop amplitudes in this interface may be considered in the future as well.

Lastly, it should be emphasised that as of today, a broad range of Monte Carlo tools with different specialisations, strengths, and weaknesses exists. In light of both, the efficiency bottlenecks and the need for revised methods to reliably calculate Monte Carlo uncertainties, it may be worthwhile to consider more streamlined methods to connect different areas of expertise and to allow for simple and meaningful scheme and model variations.

Bibliography

[1] H. G. Liddell and R. Scott, *A Greek-English lexicon*. Oxford. Clarendon Press., 1940.

[2] J. J. Thomson, *On bodies smaller than atoms*, *Popular Science Monthly* **59** (1901) .

[3] H. Geiger and E. Marsden, *On a diffuse reaction of the α particles*, *Proc.Roy.Soc.* **82** (1909) 495.

[4] H. Geiger, *The scattering of the α-particles by matter*, *Proc.Roy.Soc. A* **83** (1910) 492.

[5] E. Rutherford, *The scattering of α and β particles by matter and the structure of the atom*, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **21** (1911) 669 [https://doi.org/10.1080/14786440508637080].

[6] M. Planck, *Zur theorie des gesetzes der energieverteilung im normalspektrum*, *Verhandl. Dtsch. Phys. Ges.* **2** (1900) 237.

[7] M. Planck, *Ueber das gesetz der energieverteilung im normalspectrum*, *Annalen der Physik* **309** (1901) 553 [https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19013090310].

[8] P. Lenard, *Ueber die lichtelektrische wirkung*, *Annalen der Physik* **313** (1902) 149 [https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19023130510].

[9] A. Einstein, *Über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt*, *Annalen der Physik* **322** (1905) 132 [https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19053220607].

[10] R. A. Millikan, *A direct determination of "h."*, *Phys. Rev.* **4** (1914) 73.

[11] R. A. Millikan, *A direct photoelectric determination of planck's "h"*, *Phys. Rev.* **7** (1916) 355.

[12] N. Bohr, *On the Constitution of Atoms and Molecules*, *Phil. Mag. Ser. 6* **26** (1913) 1.

[13] A. Sommerfeld, *Zur quantentheorie der spektrallinien*, *Annalen der Physik* **356** (1916) 1 [https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19163561702].

[14] W. Heisenberg, *Über quantentheoretische umdeutung kinematischer und mechanischer beziehungen.*, *Zeitschrift für Physik* **33** (1925) 879.

[15] M. Born and P. Jordan, *Zur quantenmechanik*, *Zeitschrift für Physik* **34** (1925) 858.

[16] M. Born, W. Heisenberg and P. Jordan, *Zur quantenmechanik. ii.*, *Zeitschrift für Physik* **35** (1926) 557.

[17] E. Schrödinger, *An undulatory theory of the mechanics of atoms and molecules*, *Phys. Rev.* **28** (1926) 1049.

[18] C. D. Anderson, *The positive electron*, *Phys. Rev.* **43** (1933) 491.

[19] P. A. M. Dirac, *The Quantum Theory of the Electron*, *Proceedings of the Royal Society of London Series A* **117** (1928) 610.

[20] R. P. Feynman, *The theory of positrons*, *Phys. Rev.* **76** (1949) 749.

[21] J. Chadwick, *Possible existence of a neutron*, *Nature* **129** (1932) 312.

[22] J. Chadwick, *Bakerian lecture. – the neutron*, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **142** (1933) 1 [https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1933.0152].

[23] W. Heisenberg, *Über den Bau der Atomkerne. I*, *Zeitschrift für Physik* **77** (1932) 1.

[24] W. Heisenberg, *Über den Bau der Atomkerne. II*, *Zeitschrift für Physik* **78** (1932) 156.

[25] W. Heisenberg, *Über den Bau der Atomkerne. III*, *Zeitschrift für Physik* **80** (1933) 587.

[26] D. Iwanenko, *The Neutron Hypothesis*, *Nature* **129** (1932) 798.

[27] M. Gell-Mann, *A schematic model of baryons and mesons*, *Physics Letters* **8** (1964) 214.

[28] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking. Version 2.* 2, 1964.

[29] J. Joyce, *The Complete Novels of James Joyce*. Wordsworth Editions, 2012.

[30] CDF collaboration, *Observation of top quark production in $\bar{p}p$ collisions*, *Phys. Rev. Lett.* **74** (1995) 2626 [hep-ex/9503002].

[31] D0 collaboration, *Observation of the top quark*, *Phys. Rev. Lett.* **74** (1995) 2632 [hep-ex/9503003].

[32] W. Pauli, "Offener brief an die gruppe der radioaktiven bei der gauvereins-tagung zu tübingen." December, 1930.

[33] E. Fermi, *Versuch einer theorie der β-strahlen. i*, *Zeitschrift für Physik* **88** (1934) 161.

[34] F. Reines and C. L. Cowan jun., *The neutrino*, *Nature* **178** (1956) 446.

[35] W. Heisenberg, *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*, *Zeitschrift für Physik* **43** (1927) 172.

[36] D. P. Barber et al., *Discovery of Three Jet Events and a Test of Quantum Chromodynamics at PETRA Energies*, Phys. Rev. Lett. **43** (1979) 830.

[37] PLUTO collaboration, *Evidence for Gluon Bremsstrahlung in e+ e- Annihilations at High-Energies*, Phys. Lett. B **86** (1979) 418.

[38] TASSO collaboration, *Evidence for Planar Events in e+ e- Annihilation at High-Energies*, Phys. Lett. B **86** (1979) 243.

[39] UA1 collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at* $\sqrt{s} = 540$ *GeV*, Phys. Lett. B **122** (1983) 103.

[40] UA2 collaboration, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, Phys. Lett. B **122** (1983) 476.

[41] ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1 [1207.7214].

[42] CMS collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30 [1207.7235].

[43] CERN, *Overall view of the LHC*.

[44] C. Collaboration, *CMS collision events (7 TeV): candidate ZZ to 2e+2mu*, November, 2011.

[45] HEP Software Foundation collaboration, *A Roadmap for HEP Software and Computing R&D for the 2020s*, Comput. Softw. Big Sci. **3** (2019) 7 [1712.06982].

[46] HSF Physics Event Generator WG collaboration, *Challenges in Monte Carlo Event Generator Software for High-Luminosity LHC*, Comput. Softw. Big Sci. **5** (2021) 12 [2004.13687].

[47] ATLAS collaboration, *ATLAS HL-LHC Computing Conceptual Design Report*, .

[48] C. O. Software and Computing, *Evolution of the CMS Computing Model towards Phase-2*, tech. rep., CERN, Geneva, Jan, 2021.

[49] F. Halzen and A. D. Martin, *QUARKS AND LEPTONS: AN INTRODUCTORY COURSE IN MODERN PARTICLE PHYSICS*. 1984.

[50] G. Dissertori, I. G. Knowles and M. Schmelling, *High energy experiments and theory*. 2003.

[51] D. Griffiths, *Introduction to elementary particles*. 2008.

[52] M. Thomson, *Modern particle physics*. Cambridge University Press, New York, 2013.

[53] T. P. Cheng and L. F. Li, *GAUGE THEORY OF ELEMENTARY PARTICLE PHYSICS*. 1984.

[54] F. Mandl and G. Shaw, *QUANTUM FIELD THEORY*. 1985.

[55] S. Weinberg, *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge University Press, 6, 2005.

[56] S. Weinberg, *The quantum theory of fields. Vol. 2: Modern applications*. Cambridge University Press, 8, 2013.

[57] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.

[58] C.-N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, *Phys. Rev.* **96** (1954) 191.

[59] L. D. Faddeev and V. N. Popov, *Feynman Diagrams for the Yang-Mills Field*, *Phys. Lett. B* **25** (1967) 29.

[60] Particle Data Group collaboration, *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01.

[61] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, *Phys. Rev. Lett.* **30** (1973) 1343.

[62] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346.

[63] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and R. P. Hudson, *Experimental Test of Parity Conservation in β Decay*, *Phys. Rev.* **105** (1957) 1413.

[64] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321.

[65] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508.

[66] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585.

[67] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531.

[68] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652.

[69] B. Pontecorvo, *Inverse beta processes and nonconservation of lepton charge*, *Zh. Eksp. Teor. Fiz.* **34** (1957) 247.

[70] Z. Maki, M. Nakagawa and S. Sakata, *Remarks on the unified model of elementary particles*, *Prog. Theor. Phys.* **28** (1962) 870.

[71] J. A. Wheeler, *On the Mathematical Description of Light Nuclei by the Method of Resonating Group Structure*, *Phys. Rev.* **52** (1937) 1107.

[72] W. Heisenberg, *Die „beobachtbaren Größen“ in der Theorie der Elementarteilchen*, *Zeitschrift für Physik* **120** (1943) 513.

[73] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys. B* **126** (1977) 298.

[74] J. C. Collins and D. E. Soper, *Back-To-Back Jets in QCD*, *Nucl. Phys. B* **193** (1981) 381.

[75] J. C. Collins and D. E. Soper, *Parton Distribution and Decay Functions*, *Nucl. Phys. B* **194** (1982) 445.

[76] G. T. Bodwin, *Factorization of the Drell-Yan Cross-Section in Perturbation Theory*, *Phys. Rev. D* **31** (1985) 2616.

[77] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization for Short Distance Hadron - Hadron Scattering*, *Nucl. Phys. B* **261** (1985) 104.

[78] J. C. Collins, D. E. Soper and G. F. Sterman, *Transverse Momentum Distribution in Drell-Yan Pair and W and Z Boson Production*, *Nucl. Phys. B* **250** (1985) 199.

[79] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization of Hard Processes in QCD*, *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1 [hep-ph/0409313].

[80] R. D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244 [1207.1303].

[81] V. Bertone, S. Carrazza and J. Rojo, *APFEL: A PDF Evolution Library with QED corrections*, *Comput. Phys. Commun.* **185** (2014) 1647 [1310.1394].

[82] S. Carrazza, A. Ferrara, D. Palazzo and J. Rojo, *APFEL Web: a web-based application for the graphical visualization of parton distribution functions*, *J. Phys. G* **42** (2015) 057001 [1410.5456].

[83] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht et al., *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J. C* **75** (2015) 132 [1412.7420].

[84] A. Buckley, J. Butterworth, D. Grellscheid, H. Hoeth, L. Lonnblad, J. Monk et al., *Rivet user manual*, *Comput. Phys. Commun.* **184** (2013) 2803 [1003.0694].

[85] C. Bierlich et al., *Robust Independent Validation of Experiment and Theory: Rivet version 3*, *SciPost Phys.* **8** (2020) 026 [1912.05451].

[86] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438.

[87] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, *Sov. Phys. JETP* **46** (1977) 641.

[88] S. Catani, D. de Florian and G. Rodrigo, *Factorization violation in the multiparton collinear limit*, *PoS* **LL2012** (2012) 035 [1211.7274].

[89] A. D. Martin and M. G. Ryskin, *Higher twists in deep inelastic scattering*, *Phys. Lett. B* **431** (1998) 395 [hep-ph/9802366].

[90] A. Buckley et al., *General-purpose event generators for LHC physics*, *Phys. Rept.* **504** (2011) 145 [1101.2599].

[91] T. Sjostrand, *The Lund Monte Carlo for Jet Fragmentation*, *Comput. Phys. Commun.* **27** (1982) 243.

[92] H. U. Bengtsson, *The Lund Monte Carlo for High p_T Physics*, *Comput. Phys. Commun.* **31** (1984) 323.

[93] H. U. Bengtsson and G. Ingelman, *The Lund Monte Carlo for High p_T Physics*, *Comput. Phys. Commun.* **34** (1985) 251.

[94] T. Sjostrand, *The Lund Monte Carlo for Jet Fragmentation and e+ e- Physics: Jetset Version 6.2*, *Comput. Phys. Commun.* **39** (1986) 347.

[95] T. Sjostrand and M. Bengtsson, *The Lund Monte Carlo for Jet Fragmentation and e+ e- Physics. Jetset Version 6.3: An Update*, *Comput. Phys. Commun.* **43** (1987) 367.

[96] H.-U. Bengtsson and T. Sjostrand, *The Lund Monte Carlo for Hadronic Processes: Pythia Version 4.8*, *Comput. Phys. Commun.* **46** (1987) 43.

[97] T. Sjostrand, *High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4*, *Comput. Phys. Commun.* **82** (1994) 74.

[98] L. Lonnblad, *Development strategies for PYTHIA version 7*, *Comput. Phys. Commun.* **118** (1999) 213 [hep-ph/9810208].

[99] M. Bertini, L. Lonnblad and T. Sjostrand, *PYTHIA version 7-0.0: A Proof of concept version*, *Comput. Phys. Commun.* **134** (2001) 365 [hep-ph/0006152].

[100] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna et al., *High-energy physics event generation with PYTHIA 6.1*, *Comput. Phys. Commun.* **135** (2001) 238 [hep-ph/0010017].

[101] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852 [0710.3820].

[102] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [1410.3012].

[103] G. Marchesini and B. R. Webber, *HERWIG: A NEW MONTE CARLO EVENT GENERATOR FOR SIMULATING HADRON EMISSION REACTIONS WITH INTERFERING GLUONS*, .

[104] G. Marchesini and B. R. Webber, *HERWIG 3.0: A MONTE CARLO EVENT GENERATOR FOR SIMULATING HADRON EMISSION REACTIONS WITH INTERFERING GLUONS*, .

[105] G. Marchesini, B. R. Webber, G. Abbiendi, I. G. Knowles, M. H. Seymour and L. Stanco, *HERWIG: A Monte Carlo event generator for simulating hadron emission reactions with interfering gluons. Version 5.1 - April 1991*, *Comput. Phys. Commun.* **67** (1992) 465.

[106] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, *JHEP* **01** (2001) 010 [hep-ph/0011363].

[107] S. Gieseke, A. Ribon, M. H. Seymour, P. Stephens and B. Webber, *Herwig++ 1.0: An Event generator for e+ e- annihilation*, *JHEP* **02** (2004) 005 [hep-ph/0311208].

[108] M. Bahr et al., *Herwig++ Physics and Manual*, *Eur. Phys. J. C* **58** (2008) 639 [0803.0883].

[109] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J. C* **76** (2016) 196 [1512.01178].

[110] T. Gleisberg, S. Hoeche, F. Krauss, A. Schalicke, S. Schumann and J.-C. Winter, *SHERPA 1. alpha: A Proof of concept version*, *JHEP* **02** (2004) 056 [hep-ph/0311263].

[111] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007 [0811.4622].

[112] SHERPA collaboration, *Event Generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) 034 [1905.09127].

[113] W. Kilian, T. Ohl and J. Reuter, *WHIZARD: Simulating Multi-Particle Processes at LHC and ILC*, *Eur. Phys. J. C* **71** (2011) 1742 [0708.4233].

[114] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer and T. Stelzer, *MadGraph 5 : Going Beyond*, *JHEP* **06** (2011) 128 [1106.0522].

[115] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [1405.0301].

[116] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and collider physics*, vol. 8. Cambridge University Press, 2, 2011.

[117] J. Campbell, J. Huston and F. Krauss, *The Black Book of Quantum Chromodynamics: A Primer for the LHC Era*. Oxford University Press, 12, 2017.

[118] P. Skands, *Introduction to QCD*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Searching for New Physics at Small and Large Scales*, 7, 2012, DOI [1207.2389].

[119] S. Höche, *Introduction to parton-shower event generators*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, 11, 2014, DOI [1411.4085].

[120] S. Platzer and S. Gieseke, *Dipole Showers and Automated NLO Matching in Herwig++*, *Eur. Phys. J. C* **72** (2012) 2187 [1109.6256].

[121] C. Reuschle et al., *NLO efforts in Herwig++*, *PoS* **RADCOR2015** (2016) 050 [1601.04101].

[122] J. M. Campbell and R. K. Ellis, *An Update on vector boson pair production at hadron colliders*, *Phys. Rev. D* **60** (1999) 113006 [hep-ph/9905386].

[123] J. M. Campbell, R. K. Ellis and C. Williams, *Vector boson pair production at the LHC*, *JHEP* **07** (2011) 018 [1105.0020].

[124] J. M. Campbell, R. K. Ellis and W. T. Giele, *A Multi-Threaded Version of MCFM*, *Eur. Phys. J. C* **75** (2015) 246 [1503.06182].

[125] G. Bevilacqua, M. Czakon, M. V. Garzelli, A. van Hameren, A. Kardos, C. G. Papadopoulos et al., *HELAC-NLO*, *Comput. Phys. Commun.* **184** (2013) 986 [1110.1499].

[126] K. Arnold et al., *VBFNLO: A Parton level Monte Carlo for processes with electroweak bosons*, *Comput. Phys. Commun.* **180** (2009) 1661 [0811.4559].

[127] J. Baglio et al., *VBFNLO: A Parton Level Monte Carlo for Processes with Electroweak Bosons – Manual for Version 2.7.0*, 1107.4038.

[128] J. Baglio et al., *Release Note - VBFNLO 2.7.0*, 1404.3940.

[129] M. Chiesa, A. Denner, J.-N. Lang and M. Pellen, *An event generator for same-sign W-boson scattering at the LHC including electroweak corrections*, *Eur. Phys. J. C* **79** (2019) 788 [1906.01863].

[130] S. Actis, A. Denner, L. Hofer, A. Scharf and S. Uccirati, *Recursive generation of one-loop amplitudes in the Standard Model*, *JHEP* **04** (2013) 037 [1211.6316].

[131] S. Actis, A. Denner, L. Hofer, J.-N. Lang, A. Scharf and S. Uccirati, *RECOLA: REcursive Computation of One-Loop Amplitudes*, *Comput. Phys. Commun.* **214** (2017) 140 [1605.01090].

[132] A. Denner, J.-N. Lang and S. Uccirati, *NLO electroweak corrections in extended Higgs Sectors with RECOLA2*, *JHEP* **07** (2017) 087 [1705.06053].

[133] A. Denner, J.-N. Lang and S. Uccirati, *Recola2: REcursive Computation of One-Loop Amplitudes 2*, *Comput. Phys. Commun.* **224** (2018) 346 [1711.07388].

[134] M. Grazzini, S. Kallweit and M. Wiesemann, *Fully differential NNLO computations with MATRIX*, *Eur. Phys. J. C* **78** (2018) 537 [1711.06631].

[135] R. Boughezal, J. M. Campbell, R. K. Ellis, C. Focke, W. Giele, X. Liu et al., *Color singlet production at NNLO in MCFM*, *Eur. Phys. J. C* **77** (2017) 7 [1605.08011].

[136] J. Campbell and T. Neumann, *Precision Phenomenology with MCFM*, *JHEP* **12** (2019) 034 [1909.09117].

[137] T. Gehrmann et al., *Jet cross sections and transverse momentum distributions with NNLOJET*, *PoS* **RADCOR2017** (2018) 074 [1801.06415].

[138] A. Gehrmann-De Ridder, T. Gehrmann, N. Glover, A. Y. Huss and D. Walker, *NNLO QCD Corrections to W+jet Production in NNLOJET*, *PoS* **LL2018** (2018) 041 [1807.09113].

[139] J. Currie, A. Gehrmann-De Ridder, T. Gehrmann, N. Glover, A. Huss and J. Pires, *Jet cross sections at the LHC with NNLOJET*, *PoS* **LL2018** (2018) 001 [1807.06057].

[140] R. Gauld, N. Glover, A. Huss, I. Majer and A. Gehrmann-De Ridder, *LHC observables with NNLOJET*, *PoS* **RADCOR2019** (2019) 002.

[141] T. Becher and T. Neumann, *Fiducial q_T resummation of color-singlet processes at N^3LL+NNLO*, *JHEP* **03** (2021) 199 [2009.11437].

[142] S. Kallweit, E. Re, L. Rottoli and M. Wiesemann, *Accurate single- and double-differential resummation of colour-singlet processes with matrix+radish: w^+w^- production at the lhc*, *JHEP* **12** (2020) 147 [2004.07720].

[143] E. Boos et al., *Generic User Process Interface for Event Generators*, in *2nd Les Houches Workshop on Physics at TeV Colliders*, 9, 2001, hep-ph/0109068.

[144] J. Alwall et al., *A Standard format for Les Houches event files*, *Comput. Phys. Commun.* **176** (2007) 300 [hep-ph/0609017].

[145] S. Höche, S. Prestel and H. Schulz, *Simulation of Vector Boson Plus Many Jet Final States at the High Luminosity LHC*, *Phys. Rev. D* **100** (2019) 014024 [1905.05120].

[146] F. James, *Monte Carlo Theory and Practice*, *Rept. Prog. Phys.* **43** (1980) 1145.

[147] S. Weinzierl, *Introduction to Monte Carlo methods*, hep-ph/0006269.

[148] F. James, *Monte-Carlo phase space*, .

[149] E. Byckling and K. Kajantie, *N-particle phase space in terms of invariant momentum transfers*, *Nucl. Phys. B* **9** (1969) 568.

[150] R. Kleiss, W. J. Stirling and S. D. Ellis, *A New Monte Carlo Treatment of Multiparticle Phase Space at High-energies*, *Comput. Phys. Commun.* **40** (1986) 359.

[151] S. Plätzer, *RAMBO on diet*, 1308.2922.

[152] P. D. Draggiotis, A. van Hameren and R. Kleiss, *SARGE: An Algorithm for generating QCD antennas*, *Phys. Lett. B* **483** (2000) 124 [hep-ph/0004047].

[153] A. van Hameren and C. G. Papadopoulos, *A Hierarchical phase space generator for QCD antenna structures*, *Eur. Phys. J. C* **25** (2002) 563 [hep-ph/0204055].

[154] G. P. Lepage, *VEGAS: AN ADAPTIVE MULTIDIMENSIONAL INTEGRATION PROGRAM*, .

[155] T. Ohl, *Vegas revisited: Adaptive Monte Carlo integration beyond factorization*, *Comput. Phys. Commun.* **120** (1999) 13 [hep-ph/9806432].

[156] R. Kleiss and R. Pittau, *Weight optimization in multichannel Monte Carlo*, *Comput. Phys. Commun.* **83** (1994) 141 [hep-ph/9405257].

[157] H. Kharraziha and S. Moretti, *The Metropolis algorithm for on-shell four momentum phase space*, *Comput. Phys. Commun.* **127** (2000) 242 [hep-ph/9909313].

[158] S. Kawabata, *A New version of the multidimensional integration and event generation package BASES/SPRING*, *Comput. Phys. Commun.* **88** (1995) 309.

[159] C. G. Papadopoulos, *PHEGAS: A Phase space generator for automatic cross-section computation*, *Comput. Phys. Commun.* **137** (2001) 247 [hep-ph/0007335].

[160] P. Nason, *MINT: A Computer program for adaptive Monte Carlo integration and generation of unweighted distributions*, 0709.2085.

[161] T. Gleisberg and S. Hoeche, *Comix, a new matrix element generator*, *JHEP* **12** (2008) 039 [0808.3674].

[162] K. Kroeninger, S. Schumann and B. Willenberg, *(MC)**3 – a Multi-Channel Markov Chain Monte Carlo algorithm for phase-space sampling*, *Comput. Phys. Commun.* **186** (2015) 1 [1404.4328].

[163] S. Brass, W. Kilian and J. Reuter, *Parallel Adaptive Monte Carlo Integration with the Event Generator WHIZARD*, *Eur. Phys. J. C* **79** (2019) 344 [1811.09711].

[164] S. Weinzierl, *A General algorithm to generate unweighted events for next-to-leading order calculations in electron positron annihilation*, *JHEP* **08** (2001) 028 [hep-ph/0106146].

[165] W. T. Giele, G. C. Stavenga and J.-C. Winter, *A Forward Branching Phase-Space Generator*, 1106.5045.

[166] J. M. Campbell, W. T. Giele and C. Williams, *The Matrix Element Method at Next-to-Leading Order*, *JHEP* **11** (2012) 043 [1204.4424].

[167] T. Chen, T. M. Figy and W. T. Giele, *A Projective Phase Space Generator for Hadronic Vector Boson Plus One Jet Production*, 1907.03893.

[168] T. M. Figy and W. T. Giele, *A Forward Branching Phase Space Generator for Hadron colliders*, *JHEP* **10** (2018) 203 [1806.09678].

[169] J. Bendavid, *Efficient Monte Carlo Integration Using Boosted Decision Trees and Generative Deep Neural Networks*, 1707.00028.

[170] M. D. Klimek and M. Perelstein, *Neural Network-Based Approach to Phase Space Integration*, *SciPost Phys.* **9** (2020) 053 [1810.11509].

[171] S. Otten, S. Caron, W. de Swart, M. van Beekveld, L. Hendriks, C. van Leeuwen et al., *Event Generation and Statistical Sampling for Physics with Deep Generative Models and a Density Information Buffer*, *Nature Commun.* **12** (2021) 2985 [1901.00875].

[172] R. Di Sipio, M. Faucci Giannelli, S. Ketabchi Haghighat and S. Palazzo, *DijetGAN: A Generative-Adversarial Network Approach for the Simulation of QCD Dijet Events at the LHC*, *JHEP* **08** (2019) 110 [1903.02433].

[173] A. Butter, T. Plehn and R. Winterhalder, *How to GAN LHC Events*, *SciPost Phys.* **7** (2019) 075 [1907.03764].

[174] C. Gao, J. Isaacson and C. Krause, *i-flow: High-dimensional Integration and Sampling with Normalizing Flows*, *Mach. Learn. Sci. Tech.* **1** (2020) 045023 [2001.05486].

[175] E. Bothmann, T. Janßen, M. Knobbe, T. Schmale and S. Schumann, *Exploring phase space with Neural Importance Sampling*, *SciPost Phys.* **8** (2020) 069 [2001.05478].

[176] C. Gao, S. Höche, J. Isaacson, C. Krause and H. Schulz, *Event Generation with Normalizing Flows*, *Phys. Rev. D* **101** (2020) 076002 [2001.10028].

[177] M. L. Mangano, S. J. Parke and Z. Xu, *Duality and Multi - Gluon Scattering*, *Nucl. Phys. B* **298** (1988) 653.

[178] L. J. Dixon, *Calculating scattering amplitudes efficiently*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 95): QCD and Beyond*, 1, 1996, hep-ph/9601359.

[179] A. Kanaki and C. G. Papadopoulos, *HELAC: A Package to compute electroweak helicity amplitudes*, *Comput. Phys. Commun.* **132** (2000) 306 [hep-ph/0002082].

[180] F. Maltoni, K. Paul, T. Stelzer and S. Willenbrock, *Color Flow Decomposition of QCD Amplitudes*, *Phys. Rev. D* **67** (2003) 014026 [hep-ph/0209271].

[181] W. Kilian, T. Ohl, J. Reuter and C. Speckner, *QCD in the Color-Flow Representation*, *JHEP* **10** (2012) 022 [1206.3700].

[182] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, *JHEP* **07** (2003) 001 [hep-ph/0206293].

[183] F. Krauss, R. Kuhn and G. Soff, *AMEGIC++ 1.0: A Matrix element generator in C++*, *JHEP* **02** (2002) 044 [hep-ph/0109036].

[184] F. Maltoni and T. Stelzer, *MadEvent: Automatic event generation with MadGraph*, *JHEP* **02** (2003) 027 [hep-ph/0208156].

[185] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni et al., *MadGraph/MadEvent v4: The New Web Generation*, *JHEP* **09** (2007) 028 [0706.2334].

[186] M. Moretti, T. Ohl and J. Reuter, *O'Mega: An Optimizing matrix element generator*, hep-ph/0102195.

[187] T. Stelzer and W. F. Long, *Automatic generation of tree level helicity amplitudes*, *Comput. Phys. Commun.* **81** (1994) 357 [hep-ph/9401258].

[188] T. Gleisberg, S. Hoeche, F. Krauss and R. Matyszkiewicz, *How to calculate colourful cross sections efficiently*, 0808.3672.

[189] F. A. Berends and W. T. Giele, *Recursive Calculations for Processes with n Gluons*, *Nucl. Phys. B* **306** (1988) 759.

[190] C. Duhr, S. Hoeche and F. Maltoni, *Color-dressed recursive relations for multi-parton amplitudes*, *JHEP* **08** (2006) 062 [hep-ph/0607057].

[191] S. Catani and M. H. Seymour, *The Dipole formalism for the calculation of QCD jet cross-sections at next-to-leading order*, *Phys. Lett. B* **378** (1996) 287 [hep-ph/9602277].

[192] S. Catani and M. H. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, *Nucl. Phys. B* **485** (1997) 291 [hep-ph/9605323].

[193] D. A. Kosower, *Antenna factorization of gauge theory amplitudes*, *Phys. Rev. D* **57** (1998) 5410 [hep-ph/9710213].

[194] D. A. Kosower, *Antenna factorization in strongly ordered limits*, *Phys. Rev. D* **71** (2005) 045016 [hep-ph/0311272].

[195] J. M. Campbell and E. W. N. Glover, *Double unresolved approximations to multiparton scattering amplitudes*, *Nucl. Phys. B* **527** (1998) 264 [hep-ph/9710255].

[196] S. Catani and M. Grazzini, *Collinear factorization and splitting functions for next-to-next-to-leading order QCD calculations*, *Phys. Lett. B* **446** (1999) 143 [hep-ph/9810389].

[197] D. A. Kosower, *Multiple singular emission in gauge theories*, *Phys. Rev. D* **67** (2003) 116003 [hep-ph/0212097].

[198] D. A. Kosower, *All orders singular emission in gauge theories*, *Phys. Rev. Lett.* **91** (2003) 061602 [hep-ph/0301069].

[199] S. Catani, D. de Florian and G. Rodrigo, *The Triple collinear limit of one loop QCD amplitudes*, *Phys. Lett. B* **586** (2004) 323 [hep-ph/0312067].

[200] J. M. Campbell, *Higher order corrections to multijet production in e(+)e(-) annihilation*, Ph.D. thesis, Durham University, 1998.

[201] G. 't Hooft and M. J. G. Veltman, *Regularization and Renormalization of Gauge Fields*, *Nucl. Phys. B* **44** (1972) 189.

[202] W. T. Giele and E. W. N. Glover, *Higher order corrections to jet cross-sections in e+ e- annihilation*, *Phys. Rev. D* **46** (1992) 1980.

[203] 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) 550 [hep-ph/9401294].

[204] S. Catani, *The Singular behavior of QCD amplitudes at two loop order*, *Phys. Lett. B* **427** (1998) 161 [hep-ph/9802439].

[205] F. Buccioni, J.-N. Lang, J. M. Lindert, P. Maierhöfer, S. Pozzorini, H. Zhang et al., *OpenLoops 2*, *Eur. Phys. J. C* **79** (2019) 866 [1907.13071].

[206] V. Hirschi, R. Frederix, S. Frixione, M. V. Garzelli, F. Maltoni and R. Pittau, *Automation of one-loop QCD corrections*, *JHEP* **05** (2011) 044 [1103.0621].

[207] T. Hahn and M. Perez-Victoria, *Automatized one loop calculations in four-dimensions and D-dimensions*, *Comput. Phys. Commun.* **118** (1999) 153 [hep-ph/9807565].

[208] A. van Hameren, *OneLOop: For the evaluation of one-loop scalar functions*, *Comput. Phys. Commun.* **182** (2011) 2427 [1007.4716].

[209] R. K. Ellis and G. Zanderighi, *Scalar one-loop integrals for QCD*, *JHEP* **02** (2008) 002 [0712.1851].

[210] A. Denner, S. Dittmaier and L. Hofer, *Collier: a fortran-based Complex One-Loop LIbrary in Extended Regularizations*, *Comput. Phys. Commun.* **212** (2017) 220 [1604.06792].

[211] G. Ossola, C. G. Papadopoulos and R. Pittau, *CutTools: A Program implementing the OPP reduction method to compute one-loop amplitudes*, *JHEP* **03** (2008) 042 [0711.3596].

[212] T. Peraro, *Ninja: Automated Integrand Reduction via Laurent Expansion for One-Loop Amplitudes*, *Comput. Phys. Commun.* **185** (2014) 2771 [1403.1229].

[213] P. Mastrolia, G. Ossola, T. Reiter and F. Tramontano, *Scattering AMplitudes from Unitarity-based Reduction Algorithm at the Integrand-level*, *JHEP* **08** (2010) 080 [1006.0710].

[214] T. Binoth et al., *A Proposal for a Standard Interface between Monte Carlo Tools and One-Loop Programs*, *Comput. Phys. Commun.* **181** (2010) 1612 [1001.1307].

[215] 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) 560 [1308.3462].

[216] D. A. Kosower, *All order collinear behavior in gauge theories*, *Nucl. Phys. B* **552** (1999) 319 [hep-ph/9901201].

[217] D. A. Kosower and P. Uwer, *One loop splitting amplitudes in gauge theory*, *Nucl. Phys. B* **563** (1999) 477 [hep-ph/9903515].

[218] S. Catani and M. Grazzini, *The soft gluon current at one loop order*, *Nucl. Phys. B* **591** (2000) 435 [hep-ph/0007142].

[219] Z. Bern and D. A. Kosower, *Color decomposition of one loop amplitudes in gauge theories*, *Nucl. Phys. B* **362** (1991) 389.

[220] F. Bloch and A. Nordsieck, *Note on the Radiation Field of the electron*, *Phys. Rev.* **52** (1937) 54.

[221] T. Kinoshita, *Mass singularities of Feynman amplitudes*, *J. Math. Phys.* **3** (1962) 650.

[222] T. D. Lee and M. Nauenberg, *Degenerate Systems and Mass Singularities*, *Phys. Rev.* **133** (1964) B1549.

[223] W. T. Giele, E. W. N. Glover and D. A. Kosower, *Higher order corrections to jet cross-sections in hadron colliders*, *Nucl. Phys. B* **403** (1993) 633 [hep-ph/9302225].

[224] S. Frixione, Z. Kunszt and A. Signer, *Three jet cross-sections to next-to-leading order*, *Nucl. Phys. B* **467** (1996) 399 [hep-ph/9512328].

[225] S. Frixione, *A General approach to jet cross-sections in QCD*, *Nucl. Phys. B* **507** (1997) 295 [hep-ph/9706545].

[226] 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) 245 [hep-ph/9809429].

[227] R. Frederix, S. Frixione, F. Maltoni and T. Stelzer, *Automation of next-to-leading order computations in QCD: The FKS subtraction*, *JHEP* **10** (2009) 003 [0908.4272].

[228] J. Reuter, F. Bach, B. Chokoufe Nejad, W. Kilian, M. Stahlhofen and C. Weiss, *QCD NLO with Powheg matching and top threshold matching in WHIZARD*, *PoS* **RADCOR2015** (2016) 088 [1601.02459].

[229] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043 [1002.2581].

[230] T. Ježo and P. Nason, *On the Treatment of Resonances in Next-to-Leading Order Calculations Matched to a Parton Shower*, *JHEP* **12** (2015) 065 [1509.09071].

[231] S. Catani, S. Dittmaier, M. H. Seymour and Z. Trocsanyi, *The Dipole formalism for next-to-leading order QCD calculations with massive partons*, *Nucl. Phys. B* **627** (2002) 189 [hep-ph/0201036].

[232] R. Frederix, T. Gehrmann and N. Greiner, *Automation of the Dipole Subtraction Method in MadGraph/MadEvent*, *JHEP* **09** (2008) 122 [0808.2128].

[233] R. Frederix, T. Gehrmann and N. Greiner, *Integrated dipoles with MadDipole in the MadGraph framework*, *JHEP* **06** (2010) 086 [1004.2905].

[234] T. Gleisberg and F. Krauss, *Automating dipole subtraction for QCD NLO calculations*, *Eur. Phys. J. C* **53** (2008) 501 [0709.2881].

[235] K. Hasegawa, S. Moch and P. Uwer, *Automating dipole subtraction*, *Nucl. Phys. B Proc. Suppl.* **183** (2008) 268 [0807.3701].

[236] K. Hasegawa, S. Moch and P. Uwer, *AutoDipole: Automated generation of dipole subtraction terms*, *Comput. Phys. Commun.* **181** (2010) 1802 [0911.4371].

[237] M. H. Seymour and C. Tevlin, *TeVJet: A General framework for the calculation of jet observables in NLO QCD*, 0803.2231.

[238] M. Schönherr, *An automated subtraction of NLO EW infrared divergences*, *Eur. Phys. J. C* **78** (2018) 119 [1712.07975].

[239] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H. S. Shao and M. Zaro, *The automation of next-to-leading order electroweak calculations*, *JHEP* **07** (2018) 185 [1804.10017].

[240] R. Frederix, S. Frixione, A. S. Papanastasiou, S. Prestel and P. Torrielli, *Off-shell single-top production at NLO matched to parton showers*, *JHEP* **06** (2016) 027 [1603.01178].

[241] S. Höche, S. Liebschner and F. Siegert, *Resonance-Aware Subtraction in the Dipole Method*, *Eur. Phys. J. C* **79** (2019) 728 [1807.04348].

[242] A. Daleo, T. Gehrmann and D. Maitre, *Antenna subtraction with hadronic initial states*, *JHEP* **04** (2007) 016 [hep-ph/0612257].

[243] A. Gehrmann-De Ridder and M. Ritzmann, *NLO Antenna Subtraction with Massive Fermions*, *JHEP* **07** (2009) 041 [0904.3297].

[244] G. Abelof and A. Gehrmann-De Ridder, *Antenna subtraction for the production of heavy particles at hadron colliders*, *JHEP* **04** (2011) 063 [1102.2443].

[245] A. Gehrmann-De Ridder, T. Gehrmann and E. W. N. Glover, *Antenna subtraction at NNLO*, *JHEP* **09** (2005) 056 [hep-ph/0505111].

[246] A. Daleo, A. Gehrmann-De Ridder, T. Gehrmann and G. Luisoni, *Antenna subtraction at NNLO with hadronic initial states: initial-final configurations*, *JHEP* **01** (2010) 118 [0912.0374].

[247] G. Abelof, O. Dekkers and A. Gehrmann-De Ridder, *Antenna subtraction with massive fermions at NNLO: Double real initial-final configurations*, *JHEP* **12** (2012) 107 [1210.5059].

[248] R. Boughezal, A. Gehrmann-De Ridder and M. Ritzmann, *Antenna subtraction at NNLO with hadronic initial states: double real radiation for initial-initial configurations with two quark flavours*, *JHEP* **02** (2011) 098 [1011.6631].

[249] A. Gehrmann-De Ridder, T. Gehrmann and M. Ritzmann, *Antenna subtraction at NNLO with hadronic initial states: double real initial-initial configurations*, *JHEP* **10** (2012) 047 [1207.5779].

[250] T. Gehrmann and P. F. Monni, *Antenna subtraction at NNLO with hadronic initial states: real-virtual initial-initial configurations*, *JHEP* **12** (2011) 049 [1107.4037].

[251] A. Gehrmann-De Ridder, T. Gehrmann and E. W. N. Glover, *Infrared structure of e+ e- —> 2 jets at NNLO*, *Nucl. Phys. B* **691** (2004) 195 [hep-ph/0403057].

[252] W. Bernreuther, C. Bogner and O. Dekkers, *The real radiation antenna function for $S \to Q\bar{Q}q\bar{q}$ at NNLO QCD*, *JHEP* **06** (2011) 032 [1105.0530].

[253] W. Bernreuther, C. Bogner and O. Dekkers, *The real radiation antenna functions for $S \to Q\bar{Q}gg$ at NNLO QCD*, *JHEP* **10** (2013) 161 [1309.6887].

[254] O. Dekkers and W. Bernreuther, *The real-virtual antenna functions for $S \to Q\bar{Q}X$ at NNLO QCD*, *Phys. Lett. B* **738** (2014) 325 [1409.3124].

[255] A. Gehrmann-De Ridder, T. Gehrmann and E. W. N. Glover, *Quark-gluon antenna functions from neutralino decay*, *Phys. Lett. B* **612** (2005) 36 [hep-ph/0501291].

[256] A. Gehrmann-De Ridder, T. Gehrmann and E. W. N. Glover, *Gluon-gluon antenna functions from Higgs boson decay*, *Phys. Lett. B* **612** (2005) 49 [hep-ph/0502110].

[257] J. Currie, E. W. N. Glover and S. Wells, *Infrared Structure at NNLO Using Antenna Subtraction*, *JHEP* **04** (2013) 066 [1301.4693].

[258] T. Sjostrand and M. van Zijl, *A Multiple Interaction Model for the Event Structure in Hadron Collisions*, *Phys. Rev. D* **36** (1987) 2019.

[259] P. Z. Skands and T. Sjostrand, *Progress on multiple interactions: Modeling the underlying event in hadron hadron collisions*, *Eur. Phys. J. C* **33** (2004) S548 [hep-ph/0310315].

[260] T. Sjostrand and P. Z. Skands, *Multiple interactions and the structure of beam remnants*, *JHEP* **03** (2004) 053 [hep-ph/0402078].

[261] M. Bahr, S. Gieseke and M. H. Seymour, *Simulation of multiple partonic interactions in Herwig++*, *JHEP* **07** (2008) 076 [0803.3633].

[262] T. Sjostrand and P. Z. Skands, *Transverse-momentum-ordered showers and interleaved multiple interactions*, *Eur. Phys. J. C* **39** (2005) 129 [hep-ph/0408302].

[263] R. Corke and T. Sjostrand, *Interleaved Parton Showers and Tuning Prospects*, *JHEP* **03** (2011) 032 [1011.1759].

[264] ATLAS collaboration, *Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC*, *New J. Phys.* **13** (2011) 053033 [1012.5104].

[265] J. R. Gaunt and W. J. Stirling, *Double Parton Distributions Incorporating Perturbative QCD Evolution and Momentum and Quark Number Sum Rules*, *JHEP* **03** (2010) 005 [0910.4347].

[266] M. Diehl, D. Ostermeier and A. Schafer, *Elements of a theory for multiparton interactions in QCD*, *JHEP* **03** (2012) 089 [1111.0910].

[267] B. Blok, Y. Dokshitser, L. Frankfurt and M. Strikman, *pQCD physics of multiparton interactions*, *Eur. Phys. J. C* **72** (2012) 1963 [1106.5533].

[268] A. V. Manohar and W. J. Waalewijn, *A QCD Analysis of Double Parton Scattering: Color Correlations, Interference Effects and Evolution*, *Phys. Rev. D* **85** (2012) 114009 [1202.3794].

[269] M. Diehl, J. R. Gaunt and K. Schönwald, *Double hard scattering without double counting*, *JHEP* **06** (2017) 083 [1702.06486].

[270] B. Cabouat, J. R. Gaunt and K. Ostrolenk, *A Monte-Carlo Simulation of Double Parton Scattering*, *JHEP* **11** (2019) 061 [1906.04669].

[271] B. Cabouat and J. R. Gaunt, *Combining single and double parton scatterings in a parton shower*, *JHEP* **10** (2020) 012 [2008.01442].

[272] T. Sjostrand, *A Model for Initial State Parton Showers*, *Phys. Lett. B* **157** (1985) 321.

[273] J. R. Forshaw, J. Holguin and S. Plätzer, *Parton branching at amplitude level*, *JHEP* **08** (2019) 145 [1905.08686].

[274] J. Isaacson and S. Prestel, *Stochastically sampling color configurations*, *Phys. Rev. D* **99** (2019) 014021 [1806.10102].

[275] H. T. Li and P. Skands, *A framework for second-order parton showers*, *Phys. Lett. B* **771** (2017) 59 [1611.00013].

[276] S. Höche and S. Prestel, *Triple collinear emissions in parton showers*, *Phys. Rev. D* **96** (2017) 074017 [1705.00742].

[277] W. T. Giele, D. A. Kosower and P. Z. Skands, *Higher-Order Corrections to Timelike Jets*, *Phys. Rev. D* **84** (2011) 054003 [1102.2126].

[278] S. Platzer and M. Sjodahl, *Subleading N_c improved Parton Showers*, *JHEP* **07** (2012) 042 [1201.0260].

[279] S. Plätzer, M. Sjodahl and J. Thorén, *Color matrix element corrections for parton showers*, *JHEP* **11** (2018) 009 [1808.00332].

[280] J. Bellm, *Colour Rearrangement for Dipole Showers*, *Eur. Phys. J. C* **78** (2018) 601 [1801.06113].

[281] K. Hamilton, R. Medves, G. P. Salam, L. Scyboz and G. Soyez, *Colour and logarithmic accuracy in final-state parton showers*, 2011.10054.

[282] J. Holguin, J. R. Forshaw and S. Plätzer, *Improvements on dipole shower colour*, *Eur. Phys. J. C* **81** (2021) 364 [2011.15087].

[283] Z. Nagy and D. E. Soper, *Parton shower evolution with subleading color*, *JHEP* **06** (2012) 044 [1202.4496].

[284] Z. Nagy and D. E. Soper, *Effects of subleading color in a parton shower*, *JHEP* **07** (2015) 119 [1501.00778].

[285] S. Höche and D. Reichelt, *Numerical resummation at sub-leading color in the strongly ordered soft gluon limit*, 2001.11492.

[286] M. De Angelis, J. R. Forshaw and S. Plätzer, *Resummation and Simulation of Soft Gluon Effects beyond Leading Color*, *Phys. Rev. Lett.* **126** (2021) 112001 [2007.09648].

[287] S. Höche, D. Reichelt and F. Siegert, *Momentum conservation and unitarity in parton showers and NLL resummation*, *JHEP* **01** (2018) 118 [1711.03497].

[288] M. Dasgupta, F. A. Dreyer, K. Hamilton, P. F. Monni and G. P. Salam, *Logarithmic accuracy of parton showers: a fixed-order study*, *JHEP* **09** (2018) 033 [1805.09327].

[289] G. Bewick, S. Ferrario Ravasio, P. Richardson and M. H. Seymour, *Logarithmic accuracy of angular-ordered parton showers*, *JHEP* **04** (2020) 019 [1904.11866].

[290] M. Dasgupta, F. A. Dreyer, K. Hamilton, P. F. Monni, G. P. Salam and G. Soyez, *Parton showers beyond leading logarithmic accuracy*, *Phys. Rev. Lett.* **125** (2020) 052002 [2002.11114].

[291] J. R. Forshaw, J. Holguin and S. Plätzer, *Building a consistent parton shower*, *JHEP* **09** (2020) 014 [2003.06400].

[292] Z. Nagy and D. E. Soper, *Summations of large logarithms by parton showers*, 2011.04773.

[293] Z. Nagy and D. E. Soper, *Summations by parton showers of large logarithms in electron-positron annihilation*, 2011.04777.

[294] A. Banfi, G. P. Salam and G. Zanderighi, *Principles of general final-state resummation and automated implementation*, *JHEP* **03** (2005) 073 [hep-ph/0407286].

[295] S. Catani, L. Trentadue, G. Turnock and B. R. Webber, *Resummation of large logarithms in e+ e- event shape distributions*, *Nucl. Phys. B* **407** (1993) 3.

[296] Z. Nagy and D. E. Soper, *A parton shower based on factorization of the quantum density matrix*, *JHEP* **06** (2014) 097 [1401.6364].

[297] Z. Nagy and D. E. Soper, *What is a parton shower?*, *Phys. Rev. D* **98** (2018) 014034 [1705.08093].

[298] G. Marchesini and B. R. Webber, *Simulation of QCD Jets Including Soft Gluon Interference*, *Nucl. Phys. B* **238** (1984) 1.

[299] G. Marchesini and B. R. Webber, *Monte Carlo Simulation of General Hard Processes with Coherent QCD Radiation*, *Nucl. Phys. B* **310** (1988) 461.

[300] M. Bengtsson and T. Sjostrand, *A Comparative Study of Coherent and Noncoherent Parton Shower Evolution*, *Nucl. Phys. B* **289** (1987) 810.

[301] R. Kuhn, F. Krauss, B. Ivanyi and G. Soff, *APACIC++: A PArton Cascade In C++, version 1.0*, *Comput. Phys. Commun.* **134** (2001) 223 [hep-ph/0004270].

[302] F. Krauss, A. Schalicke and G. Soff, *APACIC++ 2.0: A Parton cascade in C++*, *Comput. Phys. Commun.* **174** (2006) 876 [hep-ph/0503087].

[303] S. Gieseke, P. Stephens and B. Webber, *New formalism for QCD parton showers*, *JHEP* **12** (2003) 045 [hep-ph/0310083].

[304] Z. Nagy and D. E. Soper, *Matching parton showers to NLO computations*, *JHEP* **10** (2005) 024 [hep-ph/0503053].

[305] Z. Nagy and D. E. Soper, *A New parton shower algorithm: Shower evolution, matching at leading and next-to-leading order level*, in *Ringberg Workshop on New Trends in HERA Physics 2005*, 1, 2006, DOI [hep-ph/0601021].

[306] S. Hoeche, S. Schumann and F. Siegert, *Hard photon production and matrix-element parton-shower merging*, *Phys. Rev. D* **81** (2010) 034026 [0912.3501].

[307] T. Carli, T. Gehrmann and S. Hoeche, *Hadronic final states in deep-inelastic scattering with Sherpa*, *Eur. Phys. J. C* **67** (2010) 73 [0912.3715].

[308] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, *JHEP* **03** (2008) 038 [0709.1027].

[309] S. Platzer and S. Gieseke, *Coherent Parton Showers with Local Recoils*, *JHEP* **01** (2011) 024 [0909.5593].

[310] S. Höche and S. Prestel, *The midpoint between dipole and parton showers*, *Eur. Phys. J. C* **75** (2015) 461 [1506.05057].

[311] G. Curci, W. Furmanski and R. Petronzio, *Evolution of Parton Densities Beyond Leading Order: The Nonsinglet Case*, *Nucl. Phys. B* **175** (1980) 27.

[312] S. Höche, F. Krauss and S. Prestel, *Implementing NLO DGLAP evolution in Parton Showers*, *JHEP* **10** (2017) 093 [1705.00982].

[313] F. Dulat, S. Höche and S. Prestel, *Leading-Color Fully Differential Two-Loop Soft Corrections to QCD Dipole Showers*, *Phys. Rev. D* **98** (2018) 074013 [1805.03757].

[314] G. Gustafson and U. Pettersson, *Dipole Formulation of QCD Cascades*, *Nucl. Phys. B* **306** (1988) 746.

[315] L. Lonnblad, *ARIADNE version 4: A Program for simulation of QCD cascades implementing the color dipole model*, *Comput. Phys. Commun.* **71** (1992) 15.

[316] W. T. Giele, D. A. Kosower and P. Z. Skands, *A simple shower and matching algorithm*, *Phys. Rev. D* **78** (2008) 014026 [0707.3652].

[317] N. Fischer, S. Prestel, M. Ritzmann and P. Skands, *Vincia for Hadron Colliders*, *Eur. Phys. J. C* **76** (2016) 589 [1605.06142].

[318] J.-C. Winter and F. Krauss, *Initial-state showering based on colour dipoles connected to incoming parton lines*, *JHEP* **07** (2008) 040 [0712.3913].

[319] B. Andersson, G. Gustafson and C. Sjogren, *Comparison of the dipole cascade model versus O(alpha-s**2) matrix elements and color interference in e+ e- annihilation*, *Nucl. Phys. B* **380** (1992) 391.

[320] G. Gustafson, *Multiplicity distributions in QCD cascades*, *Nucl. Phys. B* **392** (1993) 251.

[321] C. Friberg, G. Gustafson and J. Hakkinen, *Color connections in e+ e- annihilation*, *Nucl. Phys. B* **490** (1997) 289 [hep-ph/9604347].

[322] L. Hartgring, E. Laenen and P. Skands, *Antenna Showers with One-Loop Matrix Elements*, *JHEP* **10** (2013) 127 [1303.4974].

[323] A. J. Larkoski and M. E. Peskin, *Spin-Dependent Antenna Splitting Functions*, *Phys. Rev. D* **81** (2010) 054010 [0908.2450].

[324] A. J. Larkoski and M. E. Peskin, *Antenna Splitting Functions for Massive Particles*, *Phys. Rev. D* **84** (2011) 034034 [1106.2182].

[325] A. J. Larkoski, J. J. Lopez-Villarejo and P. Skands, *Helicity-Dependent Showers and Matching with VINCIA*, *Phys. Rev. D* **87** (2013) 054033 [1301.0933].

[326] N. Fischer, A. Lifson and P. Skands, *Helicity Antenna Showers for Hadron Colliders*, *Eur. Phys. J. C* **77** (2017) 719 [1708.01736].

[327] J. J. Lopez-Villarejo and P. Z. Skands, *Efficient Matrix-Element Matching with Sector Showers*, *JHEP* **11** (2011) 150 [1109.3608].

[328] H. Brooks and P. Skands, *Coherent showers in decays of colored resonances*, *Phys. Rev. D* **100** (2019) 076006 [1907.08980].

[329] R. Kleiss and R. Verheyen, *Final-state QED Multipole Radiation in Antenna Parton Showers*, *JHEP* **11** (2017) 182 [1709.04485].

[330] P. Skands and R. Verheyen, *Multipole photon radiation in the Vincia parton shower*, *Phys. Lett. B* **811** (2020) 135878 [2002.04939].

[331] R. Kleiss and R. Verheyen, *Collinear electroweak radiation in antenna parton showers*, *Eur. Phys. J. C* **80** (2020) 980 [2002.09248].

[332] M. Bengtsson and T. Sjostrand, *Coherent Parton Showers Versus Matrix Elements: Implications of PETRA - PEP Data*, *Phys. Lett. B* **185** (1987) 435.

[333] M. H. Seymour, *Matrix element corrections to parton shower algorithms*, *Comput. Phys. Commun.* **90** (1995) 95 [hep-ph/9410414].

[334] M. H. Seymour, *A Simple prescription for first order corrections to quark scattering and annihilation processes*, *Nucl. Phys. B* **436** (1995) 443 [hep-ph/9410244].

[335] J. Andre and T. Sjostrand, *A Matching of matrix elements and parton showers*, *Phys. Rev. D* **57** (1998) 5767 [hep-ph/9708390].

[336] N. Fischer and S. Prestel, *Combining states without scale hierarchies with ordered parton showers*, *Eur. Phys. J. C* **77** (2017) 601 [1706.06218].

[337] T. Plehn, D. Rainwater and P. Z. Skands, *Squark and gluino production with jets*, *Phys. Lett. B* **645** (2007) 217 [hep-ph/0510144].

[338] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029 [hep-ph/0204244].

[339] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [hep-ph/0409146].

[340] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070 [0709.2092].

[341] S. Hoche, F. Krauss, M. Schonherr and F. Siegert, *Automating the POWHEG method in Sherpa*, *JHEP* **04** (2011) 024 [1008.5399].

[342] S. Hoeche, F. Krauss, M. Schonherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, *JHEP* **09** (2012) 049 [1111.1220].

[343] B. Chokoufe Nejad, W. Kilian, J. Reuter and C. Weiss, *Matching NLO QCD Corrections in WHIZARD with the POWHEG scheme*, *PoS* **EPS-HEP2015** (2015) 317 [1510.02739].

[344] K. Hamilton, P. Richardson and J. Tully, *A Positive-Weight Next-to-Leading Order Monte Carlo Simulation of Drell-Yan Vector Boson Production*, *JHEP* **10** (2008) 015 [0806.0290].

[345] E. Norrbin and T. Sjostrand, *QCD radiation off heavy particles*, *Nucl. Phys. B* **603** (2001) 297 [hep-ph/0010012].

[346] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *QCD matrix elements + parton showers*, *JHEP* **11** (2001) 063 [hep-ph/0109231].

[347] K. Hamilton, P. Richardson and J. Tully, *A Modified CKKW matrix element merging approach to angular-ordered parton showers*, *JHEP* **11** (2009) 038 [0905.3072].

[348] L. Lonnblad, *Correcting the color dipole cascade model with fixed order matrix elements*, *JHEP* **05** (2002) 046 [hep-ph/0112284].

[349] L. Lonnblad and S. Prestel, *Matching Tree-Level Matrix Elements with Interleaved Showers*, *JHEP* **03** (2012) 019 [1109.4829].

[350] H. Brooks and C. T. Preuss, *Efficient multi-jet merging with the Vincia sector shower*, *Comput. Phys. Commun.* **264** (2021) 107985 [2008.09468].

[351] S. Hoeche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, *JHEP* **05** (2009) 053 [0903.1219].

[352] L. Lonnblad and S. Prestel, *Unitarising Matrix Element + Parton Shower merging*, *JHEP* **02** (2013) 094 [1211.4827].

[353] S. Plätzer, *Controlling inclusive cross sections in parton shower + matrix element merging*, *JHEP* **08** (2013) 114 [1211.5467].

[354] N. Lavesson and L. Lonnblad, *Extending CKKW-merging to One-Loop Matrix Elements*, *JHEP* **12** (2008) 070 [0811.2912].

[355] K. Hamilton and P. Nason, *Improving NLO-parton shower matched simulations with higher order matrix elements*, *JHEP* **06** (2010) 039 [1004.1764].

[356] S. Hoche, F. Krauss, M. Schonherr and F. Siegert, *NLO matrix elements and truncated showers*, *JHEP* **08** (2011) 123 [1009.1127].

[357] T. Gehrmann, S. Hoche, F. Krauss, M. Schonherr and F. Siegert, *NLO QCD matrix elements + parton showers in e^+e^- —> hadrons*, *JHEP* **01** (2013) 144 [1207.5031].

[358] S. Hoeche, F. Krauss, M. Schonherr and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, *JHEP* **04** (2013) 027 [1207.5030].

[359] L. Lönnblad and S. Prestel, *Merging Multi-leg NLO Matrix Elements with Parton Showers*, *JHEP* **03** (2013) 166 [1211.7278].

[360] J. Bellm, S. Gieseke and S. Plätzer, *Merging NLO Multi-jet Calculations with Improved Unitarization*, *Eur. Phys. J. C* **78** (2018) 244 [1705.06700].

[361] K. Hamilton, P. Nason and G. Zanderighi, *MINLO: Multi-Scale Improved NLO*, *JHEP* **10** (2012) 155 [1206.3572].

[362] R. Frederix and K. Hamilton, *Extending the MINLO method*, *JHEP* **05** (2016) 042 [1512.02663].

[363] M. L. Mangano, M. Moretti and R. Pittau, *Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b}$ + n jets as a case study*, *Nucl. Phys. B* **632** (2002) 343 [hep-ph/0108069].

[364] M. L. Mangano, M. Moretti, F. Piccinini and M. Treccani, *Matching matrix elements and shower evolution for top-quark production in hadronic collisions*, *JHEP* **01** (2007) 013 [hep-ph/0611129].

[365] R. Frederix and S. Frixione, *Merging meets matching in MC@NLO*, *JHEP* **12** (2012) 061 [1209.6215].

[366] S. Höche, Y. Li and S. Prestel, *Drell-Yan lepton pair production at NNLO QCD with parton showers*, *Phys. Rev. D* **91** (2015) 074015 [1405.3607].

[367] S. Höche, Y. Li and S. Prestel, *Higgs-boson production through gluon fusion at NNLO QCD with parton showers*, *Phys. Rev. D* **90** (2014) 054011 [1407.3773].

[368] S. Hoeche, Y. Li and S. Prestel, *Combining parton showers and NNLO matrix elements*, in *50th Rencontres de Moriond on QCD and High Energy Interactions*, 7, 2015, 1507.05325.

[369] S. Höche, S. Kuttimalai and Y. Li, *Hadronic Final States in DIS at NNLO QCD with Parton Showers*, *Phys. Rev. D* **98** (2018) 114013 [1809.04192].

[370] P. F. Monni, P. Nason, E. Re, M. Wiesemann and G. Zanderighi, *MiNNLO$_{PS}$: a new method to match NNLO QCD to parton showers*, *JHEP* **05** (2020) 143 [1908.06987].

[371] E. Re, *MiNNLO$_{\rm PS}$: a new method to match NNLO QCD with parton showers*, in *55th Rencontres de Moriond on QCD and High Energy Interactions*, 7, 2021, 2107.06753.

[372] K. Hamilton, P. Nason, E. Re and G. Zanderighi, *NNLOPS simulation of Higgs boson production*, *JHEP* **10** (2013) 222 [1309.0017].

[373] A. Karlberg, E. Re and G. Zanderighi, *NNLOPS accurate Drell-Yan production*, *JHEP* **09** (2014) 134 [1407.2940].

[374] W. Astill, W. Bizon, E. Re and G. Zanderighi, *NNLOPS accurate associated HW production*, *JHEP* **06** (2016) 154 [1603.01620].

[375] W. Astill, W. Bizoń, E. Re and G. Zanderighi, *NNLOPS accurate associated HZ production with $H \to b\bar{b}$ decay at NLO*, *JHEP* **11** (2018) 157 [1804.08141].

[376] K. Hamilton, P. Nason, C. Oleari and G. Zanderighi, *Merging H/W/Z + 0 and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching*, *JHEP* **05** (2013) 082 [1212.4504].

[377] Y. Hu, C. Sun, X.-M. Shen and J. Gao, *Hadronic decays of Higgs boson at NNLO matched with parton shower*, 2101.08916.

[378] S. Prestel, *Matching N3LO QCD calculations to parton showers*, 2106.03206.

[379] R. D. Field and R. P. Feynman, *Quark Elastic Scattering as a Source of High Transverse Momentum Mesons*, *Phys. Rev. D* **15** (1977) 2590.

[380] R. D. Field and R. P. Feynman, *A Parametrization of the Properties of Quark Jets*, *Nucl. Phys. B* **136** (1978) 1.

[381] A. Krzywicki and B. Petersson, *Breakdown of hadronic scaling or evidence for clustering?*, *Phys. Rev. D* **6** (1973) 924.

[382] P. Hoyer, P. Osland, H. G. Sander, T. F. Walsh and P. M. Zerwas, *Quantum Chromodynamics and Jets in e+ e-*, *Nucl. Phys. B* **161** (1979) 349.

[383] A. Ali, E. Pietarinen, G. Kramer and J. Willrodt, *A QCD Analysis of the High-Energy e+ e- Data from PETRA*, *Phys. Lett. B* **93** (1980) 155.

[384] B. Andersson, G. Gustafson and C. Peterson, *A Semiclassical Model for Quark Jet Fragmentation*, *Z. Phys. C* **1** (1979) 105.

[385] B. Andersson, G. Gustafson and B. Soderberg, *A General Model for Jet Fragmentation*, *Z. Phys. C* **20** (1983) 317.

[386] B. Andersson, G. Gustafson, G. Ingelman and T. Sjostrand, *Parton Fragmentation and String Dynamics*, *Phys. Rept.* **97** (1983) 31.

[387] T. Sjostrand, *Jet Fragmentation of Nearby Partons*, *Nucl. Phys. B* **248** (1984) 469.

[388] B. Andersson, *The Lund model*, vol. 7. Cambridge University Press, 7, 2005.

[389] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, *Nucl. Phys. B* **238** (1984) 492.

[390] J.-C. Winter, F. Krauss and G. Soff, *A Modified cluster hadronization model*, *Eur. Phys. J. C* **36** (2004) 381 [hep-ph/0311085].

[391] T. Sjostrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026 [hep-ph/0603175].

[392] I. G. Knowles and G. D. Lafferty, *Hadronization in Z^0 decay*, *J. Phys. G* **23** (1997) 731 [hep-ph/9705217].

[393] S. Chun and C. Buchanan, *A Simple plausible path from QCD to successful prediction of e+ e- —> hadronization data*, *Phys. Rept.* **292** (1998) 239.

[394] X. Artru and G. Mennessier, *String model and multiproduction*, *Nucl. Phys. B* **70** (1974) 93.

[395] X. Artru, *Classical String Phenomenology. 1. How Strings Work*, *Phys. Rept.* **97** (1983) 147.

[396] T. D. Gottschalk and D. A. Morris, *A New Model for Hadronization and e^+e^- Annihilation*, *Nucl. Phys. B* **288** (1987) 729.

[397] N. Isgur and J. E. Paton, *A Flux Tube Model for Hadrons*, *Phys. Lett. B* **124** (1983) 247.

[398] N. Isgur and J. E. Paton, *A Flux Tube Model for Hadrons in QCD*, *Phys. Rev. D* **31** (1985) 2910.

[399] A. Di Giacomo, M. Maggiore and S. Olejnik, *Evidence for Flux Tubes From Cooled QCD Configurations*, *Phys. Lett. B* **236** (1990) 199.

[400] A. Di Giacomo, M. Maggiore and S. Olejnik, *Confinement and Chromoelectric Flux Tubes in Lattice QCD*, *Nucl. Phys. B* **347** (1990) 441.

[401] E. Eichten, K. Gottfried, T. Kinoshita, K. D. Lane and T.-M. Yan, *Charmonium: The Model*, *Phys. Rev. D* **17** (1978) 3090.

[402] J. S. Schwinger, *On gauge invariance and vacuum polarization*, *Phys. Rev.* **82** (1951) 664.

[403] B. Andersson, G. Gustafson and T. Sjostrand, *A Model for Baryon Production in Quark and Gluon Jets*, *Nucl. Phys. B* **197** (1982) 45.

[404] B. Andersson, G. Gustafson, G. Ingelman and T. Sjostrand, *Baryon Production in Lepton - Nucleon Scattering and Diquark Fragmentation*, *Z. Phys. C* **13** (1982) 361.

[405] B. Andersson, G. Gustafson and T. Sjostrand, *Baryon Production in Jet Fragmentation and* Υ *Decay*, *Phys. Scripta* **32** (1985) 574.

[406] P. Eden and G. Gustafson, *Baryon production in the string fragmentation picture*, *Z. Phys. C* **75** (1997) 41 [hep-ph/9606454].

[407] C. Bierlich, G. Gustafson, L. Lönnblad and A. Tarasov, *Effects of Overlapping Strings in pp Collisions*, *JHEP* **03** (2015) 148 [1412.6259].

[408] J. R. Christiansen and P. Z. Skands, *String Formation Beyond Leading Colour*, *JHEP* **08** (2015) 003 [1505.01681].

[409] C. Bierlich, G. Gustafson and L. Lönnblad, *A shoving model for collectivity in hadronic collisions*, 1612.05132.

[410] N. Fischer and T. Sjöstrand, *Thermodynamical String Fragmentation*, *JHEP* **01** (2017) 140 [1610.09818].

[411] C. Bierlich, G. Gustafson and L. Lönnblad, *Collectivity without plasma in hadronic collisions*, *Phys. Lett. B* **779** (2018) 58 [1710.09725].

[412] C. B. Duncan and P. Skands, *Fragmentation of Two Repelling Lund Strings*, *SciPost Phys.* **8** (2020) 080 [1912.09639].

[413] N. Hunt-Smith and P. Skands, *String fragmentation with a time-dependent tension*, *Eur. Phys. J. C* **80** (2020) 1073 [2005.06219].

[414] G. C. Fox and S. Wolfram, *A Model for Parton Showers in QCD*, *Nucl. Phys. B* **168** (1980) 285.

[415] R. D. Field and S. Wolfram, *A QCD Model for e+ e- Annihilation*, *Nucl. Phys. B* **213** (1983) 65.

[416] Y. I. Azimov, Y. L. Dokshitzer, V. A. Khoze and S. I. Troyan, *Similarity of Parton and Hadron Spectra in QCD Jets*, *Z. Phys. C* **27** (1985) 65.

[417] Y. I. Azimov, Y. L. Dokshitzer, V. A. Khoze and S. I. Troyan, *Humpbacked QCD Plateau in Hadron Spectra*, *Z. Phys. C* **31** (1986) 213.

[418] D. Amati and G. Veneziano, *Preconfinement as a Property of Perturbative QCD*, *Phys. Lett. B* **83** (1979) 87.

[419] S. Gieseke, P. Kirchgaeßer and S. Plätzer, *Baryon production from cluster hadronisation*, *Eur. Phys. J. C* **78** (2018) 99 [1710.10906].

[420] C. B. Duncan and P. Kirchgaeßer, *Kinematic strangeness production in cluster hadronization*, *Eur. Phys. J. C* **79** (2019) 61 [1811.10336].

[421] ALICE collaboration, *Measurement of pion, kaon and proton production in proton–proton collisions at* $\sqrt{s} = 7$ *TeV*, *Eur. Phys. J. C* **75** (2015) 226 [1504.00024].

[422] L. Lonnblad, *ARCLUS: A New jet clustering algorithm inspired by the color dipole model*, *Z. Phys. C* **58** (1993) 471.

[423] H. Schulz, S. Hoeche and S. Prestel, *Z + up to 9 jets parton level events at 14 TeV in HDF5*, May, 2019. 10.5281/zenodo.2678039.

[424] H. Schulz, S. Hoeche and S. Prestel, *Wplus + up to 9 jets parton level events at 14 TeV in HDF5*, May, 2019. 10.5281/zenodo.2678055.

[425] H. Schulz, S. Hoeche and S. Prestel, *Wminus + up to 9 jets parton level events at 14 TeV in HDF5*, May, 2019. 10.5281/zenodo.2678091.

[426] C. F. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, H. Ita et al., *An Automated Implementation of On-Shell Methods for One-Loop Amplitudes*, *Phys. Rev. D* **78** (2008) 036003 [0803.4180].

[427] 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) 1981 [1209.0100].

[428] J. Reuter, B. Chokoufe, A. Hoang, W. Kilian, M. Stahlhofen, T. Teubner et al., *Automation of NLO processes and decays and POWHEG matching in WHIZARD*, *J. Phys. Conf. Ser.* **762** (2016) 012059 [1602.06270].

[429] S. Alioli, C. W. Bauer, C. Berggren, F. J. Tackmann, J. R. Walsh and S. Zuberi, *Matching Fully Differential NNLO Calculations and Parton Showers*, *JHEP* **06** (2014) 089 [1311.0286].

[430] S. Alioli, A. Broggio, A. Gavardi, S. Kallweit, M. A. Lim, R. Nagar et al., *Matching NNLO to parton shower using* N^3LL *colour-singlet transverse momentum resummation in GENEVA*, 2102.08390.

[431] A. Ballestrero et al., *Precise predictions for same-sign W-boson scattering at the LHC*, *Eur. Phys. J. C* **78** (2018) 671 [1803.07943].

[432] B. Jäger, A. Karlberg, S. Plätzer, J. Scheller and M. Zaro, *Parton-shower effects in Higgs production via Vector-Boson Fusion*, *Eur. Phys. J. C* **80** (2020) 756 [2003.12435].

[433] M. Rauch and S. Plätzer, *Parton Shower Matching Systematics in Vector-Boson-Fusion WW Production*, *Eur. Phys. J. C* **77** (2017) 293 [1605.07851].

[434] B. Jäger, F. Schissler and D. Zeppenfeld, *Parton-shower effects on Higgs boson production via vector-boson fusion in association with three jets*, *JHEP* **07** (2014) 125 [1405.6950].

[435] A. Buckley et al., *A comparative study of Higgs boson production from vector-boson fusion*, 2105.11399.

[436] ATLAS collaboration, *Modelling of the vector boson scattering process $pp \to W^{\pm}W^{\pm}jj$ in Monte Carlo generators in ATLAS*, .

[437] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, *JHEP* **02** (2010) 037 [0911.5299].