# High energy corrections to the production of Higgs-boson plus dijets in QCD

Marian Heil

A Thesis presented for the degree of
Doctor of Philosophy


Institute for Particle Physics Phenomenology
Department of Physics
Durham University
United Kingdom

September 2020

# High energy corrections to the production of Higgs-boson plus dijets in QCD

Marian Heil

Submitted for the degree of Doctor of Philosophy
September 2020

**Abstract:** With the expected increase in statistics provided by the LHC in the coming years it will enable experiments to explore a wider range of phase space then at present. Some of these regions require special treatment since the cross section induces large logarithms which lead to a breakdown of the perturbative series. In particular at large partonic energies the cross section of QCD processes is sensitive to BFKL–like logarithms of $\log|\hat{s}/\hat{t}|$. The Monte Carlo generator High Energy Jets (HEJ) resums $\log|\hat{s}/\hat{t}|$ through an expansion of the matrix element and a numerical integration of the phase space.

This thesis will develop a completely new matching of HEJ with leading-order calculations, which allows matching higher multiplicities while simultaneously decreasing the computational costs. This required a full rewrite of the source code, leading to a new, version 2 release of HEJ. The matching shown can be applied to any multi-jet process in QCD for which the matrix element in HEJ are known. In this thesis we use Higgs boson plus dijet production from gluon fusion as our first proof-of-concept for this matching.

In Higgs boson production one typically wants to isolate the weak-boson fusion production channel. Common cuts therefore require large invariant masses of the jets. Thus predictions become not only sensitive to hard emissions, but also to high-energy logarithms. Additionally even at leading order, gluon-fusion includes a massive quark loop, meaning that the process is particularly hard to accurately model. We will show how finite quark masses can be included in HEJ to all orders, and all multiplicities. The resulting simulations are some of the most accurate predictions of Higgs boson production through gluon fusion at parton colliders.

# Contents

# List of figures

# List of tables

# Declaration

The work in this thesis is based on research carried out in the Department of Physics at Durham University. No part of this thesis has been submitted elsewhere for any degree or qualification. Chapters 3 and 4 of this thesis contains joint research published as:

- Chapter 3 is based work done in collaboration with J. R. Andersen, T. Hapola, A. Maier and J. M. Smillie, published in *Higgs-boson plus Dijets: Higher-Order Matching for High-Energy Predictions*, *JHEP* **08** (2018) 090, [1805.04446] [1],
- Chapter 4 is based work done in collaboration with J. R. Andersen, J. D. Cockburn, A. Maier and J. M. Smillie, published in *Finite Quark-Mass Effects in Higgs Boson Production with Dijets at Large Energies*, *JHEP* **04** (2019) 127, [1812.08072] [2].

The High Energy Jets source code used in these papers was separately published in

- J. R. Andersen, T. Hapola, M. Heil, A. Maier and J. Smillie, *HEJ 2: High Energy Resummation for Hadron Colliders*, *Computer Physics Communications* **245** (2, 2019) 106832, [1902.08430] [3].

The newest version of the code can be obtained from `https://phab.hepforge.org/source/hej` [4] or `https://hej.web.cern.ch/HEJ/` [5].

# Acknowledgements

I would like to thank my supervisor Jeppe Andersen for his encouragement to explore different topics, even if they where not always directly related to the day-to-day research, and his guidance to still not loose focus. Without him my PhD would not have been possible, and I would not have learned as much as I did over the last three years. Further, I would like to thank Andreas Maier for all his wisdom of C++ and programming in general. I also thank all the other members of the HEJ collaboration, especially Hitham Hassan and James Black, for the many discussions.

This work has received funding from the European Union's Horizon 2020 research and innovation programme as part of the Marie Skłodowska-Curie Innovative Training Network MCnetITN3 (grant agreement no. 722104). I thank every member of MCnet for forming such an amazing community that every meeting was something special.

MCnet also allowed my to spend three month at the Lund University. I am very graceful for everyone in the theoretical physics department that welcomed me and made the stay enjoyable. Especially I would like to thank Leif Lönnblad and Stefan Prestel for explaining to me merging and parton showers in general, and Pythia in particular.

It always was a pleasure to part of the IPPP in Durham, and I am grateful for everyone that made it the place it is. In particular I thank Andrew, Christos, Francesco, Henry, Joseph, Kevin, Maria Laura and Parisa for the great office we shared and the many fun debates we had, ranging from Python and why it might not be a good programming language, over (not) learning Italian to actual physics. I also thank Alan, Duncan and Stephen for the many "successful" pub quizzes, James for proofing the integer phase space conjecture, and Robin for being honest.

Zu guter Letzt danke ich meinen Freunden und ganz besonders meiner Familie die mich erst in der Ferne und dann auch wieder ganz nah unterstützt haben. Ohne euren Rückhalt wäre das Abenteuer PhD in England so nur schwer vorstellbar.

# Chapter 1

# Introduction to the Standard Model

With the discovery of a Higgs boson in 2012 by ATLAS [6] and CMS [7] at the *Large Hadron Collider* (LHC) in CERN the venture of finding all particles in the *Standard Model* (SM) of particle physics is completed. This discovery required enormous development, experimentally and theoretically, since the proposal of the Higgs bosons in 1964 [8–10]. Prior to the LHC, the largest ever build electron-position collider—creatively named the *Large Electron–Positron Collider* (LEP)—managed to exclude masses of the Higgs boson for up to 114 GeV [11–14], just 11 GeV short of the later measured Higgs mass of 125 GeV (for a review on LEP searches see ref. 15). Moreover indirect constrained from electroweak precision data excluded Higgs boson masses above 157 GeV [16, 17]. Still LEP was not powerful enough, and instead its 27 km-tunnel got recommissioned for the LHC. After construction the LHC quickly reached record centre of mass energies of first 7 TeV in 2011 and 8 TeV in 2012, which finally lead to the discovery. In the following years the LHC got upgraded even further with an expected energy of 14 TeV in the upcoming run. Beside this massive construction task, preprocessing, storing and analysing the collected data itself also requires a huge infrastructure—both the ATLAS and CMS detector recorded an integrated luminosity of approximately $10\,\mathrm{fb}^{-1}$ alone for the Higgs discovery.[1] For example the *Worldwide LHC Computing Grid* is the world biggest computing grid, running on over 170 local clusters in 41 countries [18]. To manage all of these tasks both the ATLAS and CMS collaborations counted around 2900 members each in 2012.

Parallel to the experimental advancement, new theoretical tools were invented. Notably in *Quantum chromodynamics* (QCD) the formulation of the Dokshitzer–Gribov–

[1] The LHC hosts seven experiments, which share most of the infrastructure. Though only ATLAS and CMS are general-purpose detectors.

Lipatov–Altarelli–Parisi (DGLAP) equations [19–21] for soft and collinear emissions, and the Balitsky–Fadin–Kuraev–Lipatov (BFKL) equation [22–25] for high energy physics, gave insights into the scattering of hadrons. The first led to the formalism of *parton showers*, which are a centre piece of any general-purpose event generator (for a review see ref. 26). At least one of the "big Three"—Herwig [27], Pythia [28] and Sherpa [29]—is almost always used in any experimental study by ATLAS and CMS. Additionally, in particular regions more specialised calculations are needed for accurate simulations, e.g. at large rapidity separations between jets BFKL resummation becomes important, which are implemented in *High Energy Jets* (HEJ) [1, 3, 30–34].

More conceptually the development of the spinor-helicity formalism [35] enable a more efficient organisation of calculations, while subtraction methods allowed for cancellation of poles for numerical integration (see section 1.4.2). Which lead to *fixed-order calculations* at next-to-leading order or higher. Often these results can then be combined with parton showers, to further improve the predictions.

Even after this immense effort, the fundamental theory did not change, making the Standard Model one of the best tested theories in history. Many extensions have been proposed, but to date none was experimentally confirmed.[2] Inventing new models for unknown physics is still necessary to eventually overcome the shortcomings of the SM—like not explaining dark matter and gravity—but without experimental signals there is no guidance for these theoretical efforts. On the other hand, searches for beyond the Standard Model phenomena always require an accurate estimate of the "Standard Model background". The simulations of SM processes itself are not a "solved problem", but an active area of research. In particular after the high luminosity upgrade, the LHC aims to produce ten times more events than with the current design [38]. This will require pushing the "theoretical uncertainty" further down to match experimental statistics. Furthermore, it will be possible to explore regions in phase space that where so far at best partially accessible. Naturally the production of Higgs bosons and its interactions with other particles will be especially interesting, to verify that the discovered Higgs boson is the one predicted by the SM.

This thesis focuses on exactly such a process, which is currently not yet measured; Higgs boson production in association with two jets. An accurate prediction of this processes requires a precise modelling of QCD at large energies, which is in general not achieved in general-purpose event generators. Instead we use the resummation from HEJ, which we explain in chapter 2. To further increase the accuracy of HEJ a

[2] Depending on the chosen definition, one might or might not count neutrino oscillation through the Pontecorvo–Maki–Nakagawa–Sakata matrix [36, 37] as part of the SM. Though, this extension can neatly be incorporated into the existing theory, and at collider energies neutrinos are effectively massless.

new matching with fixed order is introduced in chapter 3. This required a full rewrite of the C++ source code of HEJ (see ref. 3–5), but allows a higher accuracy while also significantly decreasing the computational cost of simulations. Even though the new code is only presented here for Higgs boson production with jets, the general framework allows to also incorporate other QCD processes, some of which are in active development. In chapter 4 we extend HEJ to simulate Higgs plus multijet events, while keeping the full top-mass dependence in the loop induced gluon to Higgs coupling.

To begin we review the general concept of the Standard Model in this chapter. Section 1.1 starts with a brief overview of *quantum field theory* (QFT). The generation of a Higgs boson from electroweak symmetry breaking is explained in section 1.2. Afterwards we will be able to build the full SM Lagrangian for the dynamics and interaction of fields. How these fields can be measured in a particle collider is described in section 1.3. At the end of this chapter (section 1.4) we take a closer look at challenges that arise in the strong interaction. In particular, we will consider at both the high energy, ultraviolet, (section 1.4.1) and the small energy, infrared, limit (section 1.4.2). The latter is particularly important when we introduce parton showers in section 1.4.3.

## 1.1 Quantum field theory

Many equations of motion in classical field theory have an analogous equation in quantum field theory. For example the classical *Euler-Lagrange equations*

$$\frac{\partial L(x_i, {}^{\mathrm{d}x_i}/_{\mathrm{d}t})}{\partial x_i} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L(x_i, {}^{\mathrm{d}x_i}/_{\mathrm{d}t})}{\partial\left({}^{\mathrm{d}x_i}/_{\mathrm{d}t}\right)}\right) \tag{1.1.1}$$

describes the motion of some objects at $x_i$ given by the *Lagrangian* $L$. Equivalently

$$\frac{\partial \mathcal{L}(\phi_i, \partial_\mu \phi_i)}{\partial \phi_i} = \partial_\mu \left(\frac{\partial \mathcal{L}(\phi_i, \partial_\mu \phi_i)}{\partial\left(\partial_\mu \phi_i\right)}\right), \tag{1.1.2}$$

is the evolution of fields $\phi_i(x)$ in four dimensions $\mu = 0, \ldots 3$. Notably the Lagrangian $\mathcal{L}$[3] itself is also a field. In both cases the Euler-Lagrange equations are derived by extremising the action

$$S \equiv \int \mathrm{d}^4x\, \mathcal{L}(\phi_i(x), \partial_\mu \phi_i(x)). \tag{1.1.3}$$

Here and in the following we choose natural units $\hbar = c = 1$.

[3]Technically $\mathcal{L}$ is the *Lagrangian density*. We will use *Lagrangian* $L$ and *Lagrangian density* $\mathcal{L}$ interchangeably.

The main difference lies in the interpretation of $x_i$ and $\phi_i$. In a classical system $x_i$ is a measurable property, like the position of some object in space. In QFT however, operators, $\phi_i$, $\mathcal{L}$, etc., cannot be measured directly. Instead they transform a state $|s\rangle$ into another state $|s'\rangle = \phi\,|s\rangle$. Measurements are eigenvalues of Hermitian operators $\phi$, i.e. the *expectation value* of finding the original state $s$ after applying $\phi$ is

$$\left|\langle s|s'\rangle\right|^2 = |\langle s|\phi|s\rangle|^2\,. \tag{1.1.4}$$

We further want to construct a relativistic theory, and thus the Lagrangian to be invariant under *Poincaré symmetry*, equivalent to invariance under Galilean transformation in classical mechanics. It follows from Noether's theorem that Poincaré invariance implies the conservation of four-momentum and angular momentum. The invariant Lagrangian for spin-$\frac{1}{2}$ particles $\psi$, *fermions*, is then

$$\mathcal{L}_{\text{fermi}} = \bar{\psi}(x)\left(\mathrm{i}\partial\!\!\!/ - m\right)\psi(x), \tag{1.1.5}$$

where we used the short-hand *Feynman slash notation* $\partial\!\!\!/ \equiv \gamma^\mu\partial_\mu$ and conjugated field

$$\bar{\psi}(x) = \psi^\dagger\gamma^0. \tag{1.1.6}$$

The $\gamma^\mu$ matrices fulfil the *Clifford algebra*

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbb{1}_4. \tag{1.1.7}$$

The fermion $\psi$ transforms like a *spinor* under *Lorentz transformations* $\Lambda$. Thus, eq. (1.1.5) is an equation in spinor space, where each term forms a *Lorentz scalar*.

Applying the Euler-Lagrange equation (1.1.2) to the conjugated field yields the *Dirac equation*

$$\left(\mathrm{i}\partial\!\!\!/ - m\right)\psi(x) = 0, \tag{1.1.8}$$

which is solved by $\psi(x) = u(p)\mathrm{e}^{-\mathrm{i}p\cdot x}$ with the four-momentum $p^\mu = (E, \vec{p})$. Further for the conjugated field we find a second, linearly independent solution $\bar{\psi}(x) = v(p)\mathrm{e}^{+\mathrm{i}p\cdot x}$. The *particle* solution $u$ propagates forward in time, while the *antiparticle* $v$ propagates backwards.

Equation (1.1.7) is not enough to uniquely define $\gamma^\mu$, thus for explicitly solving eq. (1.1.8) we have to choose a representation. We will use the *Weyl basis* in which

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}, \tag{1.1.9}$$

with $\sigma^\mu = (\mathbb{1}, \vec{\sigma})$, $\bar{\sigma}^\mu = (\mathbb{1}, -\vec{\sigma})$, and $\sigma^i$ are the *Pauli matrices*. In this representation eq. (1.1.8) simplifies to

$$\left(\not{p} - m\right) u(p) = 0, \tag{1.1.10}$$

$$\left(\not{p} + m\right) v(p) = 0, \tag{1.1.11}$$

giving two solutions each

$$u_s(p) = \begin{pmatrix} \sqrt{p \cdot \sigma}\, \xi_s \\ \sqrt{p \cdot \bar{\sigma}}\, \xi_s \end{pmatrix}, \tag{1.1.12}$$

$$v_s(p) = \begin{pmatrix} \sqrt{p \cdot \sigma}\, \eta_s \\ -\sqrt{p \cdot \bar{\sigma}}\, \eta_s \end{pmatrix}. \tag{1.1.13}$$

We are using the notation from ref. 39, where the root $\sqrt{p \cdot \sigma}$ implies diagonalising $p \cdot \sigma$, such that we can take the root of each component, e.g. with $p^\mu$ aligned along the $z$-axis

$$\sqrt{p \cdot \sigma} = \begin{pmatrix} \sqrt{E - p_z} & 0 \\ 0 & \sqrt{E + p_z} \end{pmatrix}. \tag{1.1.14}$$

$\xi_s$ and $\eta_s$ are unit basis vectors in the two-component spinor space, where $s = 1, 2$ is the *spin* of the fermions. We can further define the *chirality* through the projection

$$P_{\mathrm{L/R}} = \frac{1 \mp \gamma_5}{2}, \tag{1.1.15}$$

with

$$\gamma_5 = \mathrm{i}\gamma_0\gamma_1\gamma_2\gamma_3 = \begin{pmatrix} -\mathbb{1}_2 & 0 \\ 0 & \mathbb{1}_2 \end{pmatrix}. \tag{1.1.16}$$

Thus each spinor decomposes into two *Weyl*-spinors, or, equivalently, their left- and right-handed components $u = (u_\mathrm{L}\ u_\mathrm{R})^T$.

The Standard Model is a QFT based on Yang-Mills theory; each force corresponds to an (local) gauge symmetry SU($N$), which at once transforms $N$ fermions

$$\vec{\psi} \to U\vec{\psi}, \tag{1.1.17}$$

where $U = \exp\left(\mathrm{i}\alpha_a T^a\right)$ is an unitary $N \times N$ matrix, and $T^a$ are the generators of the symmetry group. However, the Lagrangian eq. (1.1.5) is only invariant under this transformation

$$\mathcal{L} \to \mathcal{L}' = \bar{\psi}_i U_{ij}^{-1} \left(\mathrm{i}\partial\!\!\!/ - m\right) U_{jk}\psi_k \tag{1.1.18}$$

$$= \bar{\psi}_i \left(\mathrm{i}\partial\!\!\!/ - m\right) \psi_i + \mathrm{i}\bar{\psi}_i U_{ij}^{-1} \partial\!\!\!/ U_{jk}\psi_k\,, \tag{1.1.19}$$

if $U$ is a *global symmetry*.

To restore local gauge symmetry, we need to modify the Lagrangian for fermions, such that it cancels the last term from eq. (1.1.19). Thus, we replace the partial derivative $\partial_\mu$ by the *covariant derivative* $D_\mu$

$$\partial_\mu \to D_\mu = \partial_\mu - \mathrm{i}gA^a_\mu(x)T^a, \tag{1.1.20}$$

with the *coupling strength* $g$ being a free parameter. We further introduce $N^2 - 1$ new fields $A^a_\mu$, which transform like vectors under Lorentz transformation. Thus we have added new bosons, call *gauge bosons*, to our theory. The resulting covariant derivative transforms like a field $\psi$

$$D_\mu\psi \to U \cdot D_\mu\psi \tag{1.1.21}$$

$$\Leftrightarrow \quad A^a_\mu T^a \to U A^a_\mu T^a U^{-1} - \frac{i}{g}\left(\partial_\mu U\right) U^{-1}. \tag{1.1.22}$$

By construction the fermionic Lagrangian eq. (1.1.5) is therefore invariant under local gauge symmetry.

Plugging the covariant derivative eq. (1.1.20) into the fermionic Lagrangian eq. (1.1.5) gives a new term

$$\mathcal{L}_{\text{int}} = gT^a_{ij}\bar{\psi}_i \not{A}^a \psi_j \,, \tag{1.1.23}$$

in which the gauge bosons $A_\mu$ couple to the fermions $\psi_i$ and $\psi_j$. $T^a_{ij}$ acts as the *charge* of the fermions. Since the new term is strictly necessary by gauge transformation, the interaction of fermions with the gauge boson is fundamental in the theory. Thus one can produce bosons.

So far, the Euler-Lagrange equation for $A_\mu$ from eq. (1.1.23) implies $\bar{\psi}\psi = 0$. To remedy this, we have to add kinematic terms for the bosons. We construct the *field strength tensors* $F_{\mu\nu}$ from the commutator of the covariant derivatives

$$F^a_{\mu\nu}T^a = \frac{\mathrm{i}}{g}\left[D_\mu, D_\nu\right] \tag{1.1.24}$$

$$\Rightarrow \qquad F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu \,, \tag{1.1.25}$$

where the structure constants of the generators $f^{abc}$ are defined as

$$\left[T^a, T^b\right] = \mathrm{i}f^{abc}T^c. \tag{1.1.26}$$

For abelian groups the generators commute, and $f^{abc} = 0$.

The Lorentz contraction of $F^a_{\mu\nu}$ with itself lead to a gauge invariant Lagrangian[4]

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} F^a_{\mu\nu} F^{a,\mu\nu} \tag{1.1.27}$$

$$= -\frac{1}{4}\left(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu\right)^2 - g f^{abc}(\partial_\mu A^a_\nu) A^{b,\mu} A^{c,\nu} - \frac{g^2}{4} f^{abc} A^b_\mu A^c_\nu f^{ade} A^{d,\mu} A^{e,\nu} . \tag{1.1.28}$$

The first term describes the kinematic of massless bosons, while the last two terms, proportional to $f^{abc}$, only appear in non-abelian gauge symmetries. For the abelian U(1) symmetry, the resulting Euler-Lagrange equations in the Lorenz gauge $\partial_\mu A^\mu = 0$ are wave equations, similar to classical electrodynamics. Contrary to this, for non-abelian gauge symmetries these extra terms correspond to gauge boson self-interaction, between three or four bosons.

The Standard Model of particle physics has three fundamental symmetries, $\mathrm{U}(1)_Y \otimes \mathrm{SU}(2)_L \otimes \mathrm{SU}(3)_C$, leading to 12 gauge bosons; one boson $B_\mu$ from $\mathrm{U}(1)_Y$, three bosons $W^a_\mu$ from SU(2) and eight gluons from SU(3). Before electroweak symmetry breaking (see next section), the charge of $\mathrm{U}(1)_Y$ is the hypercharge, while SU(2) has two weak isospins. Quantum chromodynamics $\mathrm{SU}(3)_C$ has three charges, called colours— red, green and blue—in analogy to the three primary colours.

## 1.2 Higgs mechanism

In the previous section we introduced gauge bosons to restore local gauge invariance. To avoid breaking gauge symmetry, the resulting bosons had to be massless. However, experimentally we know that in nature the weak bosons are massive; the electroweak symmetry has to be broken through the so-called *Higgs mechanism* [8–10]. We start by introducing a new complex SU(2) doublet, the *Higgs multiplet*,

$$h = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + \mathrm{i}\phi_2 \\ \phi_3 + \mathrm{i}\phi_4 \end{pmatrix}, \tag{1.2.1}$$

with hypercharge $Y = \frac{1}{2}$, and four degrees of freedom corresponding to the four real scalars $\phi_i$. The Lagrangian

$$\mathcal{L}_h = \left(D_\mu h\right)^\dagger D^\mu h + \mu^2 h^\dagger h - \lambda\left(h^\dagger h\right)^2, \tag{1.2.2}$$

[4]There is a second gauge invariant term $\epsilon^{\mu\nu\xi o} F^a_{\mu\nu} F^a_{\xi o} = 2\epsilon^{\mu\nu\xi o}\partial_\mu\left(A^a_\nu F^a_{\xi o}\right)$, which is a total derivative. Hence, this term changes the action only on the boundaries, without affecting the equations of motions. So it does not contribute in perturbative theories.

is then invariant under $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ when using the corresponding covariant derivative

$$D_\mu = \partial_\mu - ig\, W_\mu^a \tau^a - \frac{1}{2} \mathrm{i} g' B_\mu \,. \tag{1.2.3}$$

$\mu$ and $\lambda$ are free parameters of the Higgs potential eq. (1.2.2). If $\mu^2 > 0$ the minimum of the Higgs potential is at $|h| \neq 0$. We arbitrary choose the minimum to be in the second component, such that the *vacuum expectation value* is

$$\langle h \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{1.2.4}$$

where

$$v = \sqrt{\frac{\mu^2}{\lambda}} \,. \tag{1.2.5}$$

We can expand $h$ around eq. (1.2.4). In the unitary gauge[5] all but one real scalar—i.e. all the Goldstone bosons—vanish, such that

$$h = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} . \tag{1.2.6}$$

Thus the Higgs mechanism leads to one new scalar boson $H$, the *Higgs boson*.

Substituting eq. (1.2.6) into the kinetic term of the Higgs multiplet

$$\begin{aligned} \left(D_\mu h\right)^\dagger D^\mu h = \frac{1}{2} \left(\partial_\mu H\right)^2 + \frac{1}{8} (v+H)^2 \Big[ & g^2 \left( W_\mu^1 + \mathrm{i}\, W_\mu^2 \right) \left( W_\mu^1 - \mathrm{i}\, W_\mu^2 \right) \\ & + \left( g\, W_\mu^3 - g' B_\mu \right)^2 \Big] , \end{aligned} \tag{1.2.7}$$

give rise to terms proportional to $v^2$, which almost look like boson masses, but with a non-diagonal mass matrix. Thus we have to linearly combine different bosons to diagonalise the masses

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left( W_\mu^1 \mp W_\mu^2 \right), \quad Z_\mu = \cos\theta_W\, W_\mu^3 - \sin\theta_W B_\mu \,. \tag{1.2.8}$$

Each of these new bosons then have their respective masses

$$m_W = g \frac{v}{2}, \quad m_Z = \sqrt{g^2 + g'^2}\, \frac{v}{2} = \frac{m_W}{\cos\theta_W} \,. \tag{1.2.9}$$

Consequently, the $Z$ boson has to be heavier then the $W^\pm$ bosons, where the ratio

[5] For the corresponding equations without gauge fixing see, for example, ref. 40.

| Is coloured? | Electric charge | Generation | | |
|---|---|---|---|---|
| Yes, quarks | $+2/3$ | up | charm | top |
| | $-1/3$ | down | strange | bottom |
| No, leptons | $-1$ | $e^-$ | $\mu^-$ | $\tau^-$ |
| | 0 | $\nu_e$ | $\nu_\mu$ | $\nu_\tau$ |

Table 1.1: Fermion content of the Standard Model.

of their masses is the *Weinberg angle*[6]

$$\tan\theta_W = \frac{g'}{g}\,. \tag{1.2.10}$$

Since the mass eigenstates are not identical to the gauge eigenstates the electroweak symmetry is broken.

The rotation in $\mathrm{U}(1)_Y \times \mathrm{SU}(2)_L$ orthogonal to the $Z$ boson is the massless photon

$$A_\mu = \sin\theta_W\, W^3_\mu + \cos\theta_W B_\mu\,. \tag{1.2.11}$$

It is the gauge boson of *quantum electrodynamics* (QED), an unbroken electromagnetic symmetry $\mathrm{U}(1)_{\mathrm{EM}}$ with coupling strength $e = g'\cos\theta_W$. The photon directly interacts with the $W^\pm$ bosons, which have an electromagnetic charge of $\pm 1$ (hence their names).

The remaining terms of eq. (1.2.7) describe the kinematics of the Higgs boson, and the coupling of up to two Higgs bosons with two $Z$ or two $W^\pm$ bosons respectively. The Higgs mass follows from eq. (1.2.2)

$$m_H = \sqrt{2}\mu \approx 125\,\mathrm{GeV}. \tag{1.2.12}$$

To couple the Higgs boson to fermions we introduce the *Yukawa couplings*, i.e. the left-handed SU(2) doublet

$$L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \tag{1.2.13}$$

interacts with the right handed singlet $e_R$ and a Higgs boson

$$\mathcal{L}_{\mathrm{Yuk}} = -y_e \bar{L} h e_R + \mathrm{h.c.} \tag{1.2.14}$$

[6] More generally eq. (1.2.9) is a consequence of a residual, global SU(2) symmetry, called *custodial symmetry* [41, 42]. After radiative corrections the mass ratio derivatives slightly from eq. (1.2.9), which can be used to constrain the free parameters of the standard model through electroweak precision data.

Similarly to the bosons, the Higgs mechanism then generates the fermion masses[7]

$$m_e = \frac{y_e}{\sqrt{2}} v. \tag{1.2.15}$$

The coupling strength of the Higgs boson to fermion interaction is therefore proportional to the fermion mass. The particle content of the SM after electroweak symmetry breaking is summarised in table 1.1.

With this we constructed all the pieces required for the Lagrangian of the Standard Model. Each local gauge symmetry required the introduction of new gauge bosons. These bosons then mediate their respective forces, either strong, electromagnetic or weak. Through electroweak symmetry breaking we generated mass terms for the weak bosons (and fermions). Thereby this also introduced a new scalar particle, the Higgs boson, with a new Yukawa interaction.

## 1.3 Scattering cross section

So far we derived terms for the kinematics and interaction of fields but, as stated earlier, we can not measure fields directly. Instead we measure the expectation value for the interference between states (cf. eq. (1.1.4)). In the Heisenberg picture that means we are interested in the transition operator $S$, which transforms an (asymptotic) incoming to an (asymptotic) outgoing state

$$\langle \text{out}|\text{in}\rangle = \lim_{t\to\infty} \left\langle k_i \middle| S(t) \middle| p_j \right\rangle, \tag{1.3.1}$$

$$\tag{1.3.2}$$

$p_i$ and $k_j$ are the states of the incoming and outgoing particles respectively. The $S$-matrix is unitary ($S^\dagger S = \mathbb{1}$), we can split off its trivial part

$$S = \mathbb{1} + \mathrm{i}\mathcal{T}. \tag{1.3.3}$$

The remaining *transfer matrix* $\mathcal{T}$ includes all state changes. We further factor out energy-momentum conservation

$$\mathcal{T} = (2\pi)^4 \delta^4\Big(\sum_i p_i^\mu + \sum_j k_j^\mu\Big)\mathcal{M}, \tag{1.3.4}$$

to end up with the (Lorentz invariant) *amplitude* $\mathcal{M}$.

To compactly write down amplitudes in perturbative QFT one can use *Feynman*

[7] In general the mass and flavour basis of fermions are different as well. Up- and down-type quarks mix via the *Cabibbo–Kobayashi–Maskawa* [43, 44] matrix when interacting with a $W^\pm$.

*diagrams*, which express complex mathematical terms through symbolic diagrams, with simple construction rules. Each interaction of the SM Lagrangian acts as an *vertex* connecting different particles. These connections come from the kinematics of the particles. Internal lines describe virtual fields, that only *propagate* an interaction, while external lines are real, on-shell emissions. A complete set of Feynman rules for all possible particles of the SM can be found in any related textbook—e.g. ref. 39, 40, 42—and will not be repeated here.

Scatterings are measured through the *cross section*

$$\sigma = \frac{\text{number of scattered particles}/\text{time}}{\text{density of particles in overlapping beams}/\text{time}}, \tag{1.3.5}$$

which, geometrically, corresponds to an area of interaction. We can directly relate the cross section to the transition probability in QFT, and therefore to the amplitude. For a $2 \to n$ scattering process at partonic invariant mass $\hat{s}$ in a collider experiment this gives the cross section

$$\begin{aligned}\sigma(\mathcal{O}) = \int \left( \prod_j \frac{\mathrm{d}^3 \vec{k}_j}{(2\pi)^3} \frac{1}{2E_j} \right) (2\pi)^4 \delta^4(p_a^\mu + p_b^\mu + \sum_j k_j^\mu) \frac{\mathcal{F}(p_a, p_b)}{\hat{s}^2} \\ \mathcal{O}(\{k_j\}) \left| \mathcal{M}(p_a p_b \to \{k_j\}) \right|^2 ,\end{aligned} \tag{1.3.6}$$

which depends on an arbitrary observable $\mathcal{O}(\{k_j\})$. For example $\mathcal{O}$ can include specific cuts required by the experiment, or by specific measurements. The total cross section $\sigma$ follows from the fully inclusive $\mathcal{O} = 1$. Here $p_a$ and $p_b$ are the incoming, and $k_j$ the outgoing momenta respectively. The first term in eq. (1.3.6) is the *Lorentz-invariant phase space*, followed by a $\delta$ function for momentum conservation and a *flux* $\mathcal{F}$. The flux is the likelihood of finding the incoming particles inside the beam. Thus calculations in QFT break down to calculating the amplitude square $|\mathcal{M}|^2$, also called *matrix element*, and integrating it over the desired region.

For example in a fixed-order calculations one expands $\mathcal{M}$ in the (strong) coupling strength $\alpha_s$, to get a cross section

$$\mathrm{d}\sigma^{\mathrm{N}^n\mathrm{LO}} = \sum_{i=0}^{n} \alpha_s^{l+i} \mathrm{d}\sigma^i, \tag{1.3.7}$$

where $l$ is the minimal order at which the process is possible, and $\sigma^i$ is the cross section at order $n$. For a long time in QCD only leading order (LO), or *Born-level*, calculations were available for most processes, but with the development of general subtraction methods (e.g. [45–47]), the automation of loop calculations [48–53], and the increase in computing power next-to-leading order (NLO) became the de-facto standard [29, 54, 55]. More recently, some processes with up to five external particles

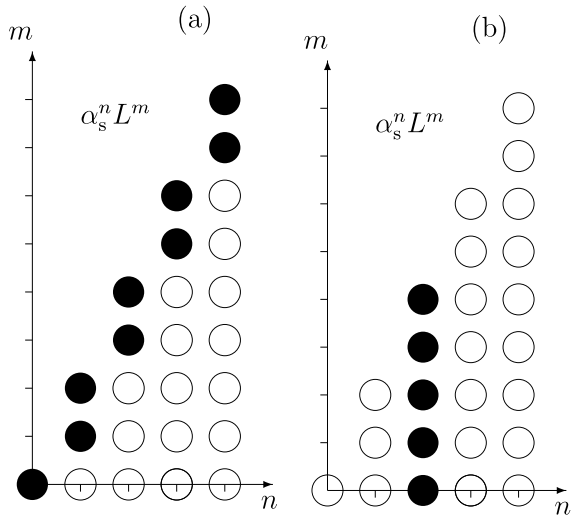



Figure 1.1: Visualisation of (a) resummation and (b) fixed order. Each dot represents one term in the all-order matrix element. Full dots are included, while empty ones are missing in the calculation. $n$ and $m = n - l$ are the powers of the strong coupling $\alpha_s$ and the logarithm respectively. The resummation is shown up to next-to-leading logarithmic accuracy $l = 1$. From ref. 26.

where computed at $N^2LO$ or even $N^3LO$ (for an overview see the list in ref. 56).

Expanding in $\alpha_s$ will break down if some common logarithm $L = \log x$ for any observable $x$ in all $d\sigma^i$ becomes large, i.e. $\alpha_s \log x \sim 1$ for all orders $i$. These logarithms then have to be *resummed*; instead of expanding in $\alpha_s$ one sums $\alpha_s^n \log^{n-l} x$ for all $n$ at a fixed (logarithmic) order $l$. A resummation therefore includes infinitely many powers of $\alpha_s$ (see fig. 1.1). As a downside resummations are limited to their respective regions of phase space where $\log x \gg 1$, which might not be dominant or even accessible in experiments. Most commonly resummations are used in the soft and collinear limit through parton showers, which is briefly reviewed in section 1.4.3. Section 2.3 introduces a different framework to resum high energy logarithms, which is basis for the rest this thesis.

To improve resummations away from their strict limits one can match and merge them with fixed-order calculations. The first few orders in $\alpha_s$ are therefore taken from fixed order. Afterwards resummation describes the higher $\alpha_s$ order terms. This will be explicitly shown for the high energy region in chapter 3.

## 1.4 Quantum chromodynamics

The strong force is the dominant interaction for hadrons, and thus also dominates LHC processes. As such a precise modelling of Quantum Chromodynamics is required

for almost every measurement from ATLAS and CMS—either to measure QCD directly or to estimate the QCD background. QCD is also the only non-abelian, unbroken gauge symmetry in the Standard Model. In this section we will highlight two key features; *asymptotic freedom*, coming from the running of the strong coupling, and infrared (IR) divergences with *confinement*, leading to parton showers. QCD describes the interaction of *coloured* particles; namely the interaction between the gauge bosons, called gluons, and a subset of all SM fermions, called quarks. Unlike photons in QED, gluons carry two (colour) charges, thus they interact directly with each other. Consequently, keeping track of the correct charges is much more complicated in QCD compared to QED. It is therefore often useful to (almost) independently consider *colour flow* and the kinematics (see appendix A).

We begin with definitions of common factors, which we will use throughout this thesis. As we saw before, QCD has $N_C = 3$ colours leading to $N_C^2 - 1 = 8$ generators $t^a$ of SU(3), and hence 8 gluons. These generators fulfil the usual commutation relations eq. (1.1.26) with the conventional normalisation

$$T^a_{ij} T^b_{ji} = \delta^{ab} T_R \equiv \delta^{ab} \frac{1}{2} \,. \tag{1.4.1}$$

Often an explicit representation for $T^a$ is unnecessary, since the common terms reappear regularly; most prominently the Casimirs, also called *colour factors*

$$T^a_{ij} T^a_{jk} = \delta_{ik} C_F = \delta_{ik} \frac{N_C^2 - 1}{2N_C} \,, \tag{1.4.2}$$

$$f^{abc} f^{abd} = \delta^{cd} C_A = \delta^{cd} N_C \tag{1.4.3}$$

of the fundamental and adjoint group representation respectively. For example summing over all colours in quark-gluon vertices gives $C_F$, while pure gluon vertices lead to $C_A$. We further assume that most quarks are light enough to ignore their masses, thus the number of light quarks, $n_f$, is 4 or 5, dependent on whether we consider both top- and bottom-quarks or only the top-quark to be massive. Together with the gluon, the light quarks are sometimes referred to as *partons*.

### 1.4.1 Renormalisation & running coupling

When treated perturbatively, QFTs inherently experience corrections from higher order loop diagrams. These loops lead to high energy, *ultraviolet* (UV) divergences, which have to be regularised. Any observable $\mathcal{O}$ then picks up a dependency on an unphysical *renormalisation scale* $\mu_r$. Measurable observables have to be independent

of this scale, which is expressed through the *Renormalisation Group Equations*

$$0 = \mu_r^2 \frac{\mathrm{d}}{\mathrm{d}\mu_r^2} \mathcal{O}\left(\log \frac{q^2}{\mu_r^2}, \alpha_s(\mu_r^2)\right) \tag{1.4.4}$$

$$= \left[\frac{\partial}{\partial \log(q^2/\mu_r^2)} + \beta(\alpha_s)\partial_{\alpha_s}\right] \mathcal{O}\left(\log \frac{q^2}{\mu_r^2}, \alpha_s(\mu_r^2)\right), \tag{1.4.5}$$

where

$$\beta(\alpha_s) = \mu_r^2 \frac{\partial \alpha_s(\mu_r^2)}{\partial \mu_r^2}, \tag{1.4.6}$$

and $\alpha_s = g_s^2/4\pi$ is the strong coupling constant. The $\mu_r$-dependence is therefore moved from the observable to the coupling, and—not shown here—the particle masses. As such a running coupling is predicted from the QFT alone, without external input. It can even be measured to test the SM at high precision over many different scales [57].

We expand eq. (1.4.6) around 0 for $\alpha_s \ll 1$

$$\beta(\alpha_s) = -\frac{\alpha_s^2}{4\pi} \sum_{i=0} \beta_i \left(\frac{\alpha_s}{4\pi}\right)^i. \tag{1.4.7}$$

At leading order this is solved analytically

$$\alpha_s(\mu_r^2) = \frac{2\pi}{\beta_0 \ln \frac{\mu_r^2}{\Lambda^2}}. \tag{1.4.8}$$

For QCD the $\beta$-function is know for up to five-loops [58–60], with the leading term

$$\beta_0 = \frac{11C_A - 4T_R n_f}{3} > 0. \tag{1.4.9}$$

Hence with increasing scale the coupling decreases; QCD is *asymptotically free*. Only at infinite energies can a quark or gluon be observed as a free particle. At small energies they always interact with each other, making a direct observation of coloured final states impossible. Contrary in QED, $\beta_0^{\mathrm{QED}} < 0$, this leads to the electric charge vanishing in the low energy, *infrared* limit. The weak interaction on the other hand also has $\beta_0 > 0$, but it is suppressed below the $W^\pm$- and $Z$-boson mass thresholds. Consequently, soft physics is dominated by the strong force.

For $\mu_r = \Lambda = \mathcal{O}(0.2\,\mathrm{GeV})$ the strong coupling diverges. This coincides with *colour-confinement*; partons always group into colour-neutral (white) hadrons. Importantly $\Lambda$ sets a lower bound on perturbative calculations. As one gets close to $\alpha_s \sim 1$ other methods have to be used, e.g. lattice QCD. In perturbative QCD the cross section factorises into one perturbative and one non-perturbative part, i.e. a factorisation of large and small energy effects respectively.

In the latter case we model a hadron as a collection of many particles. Each

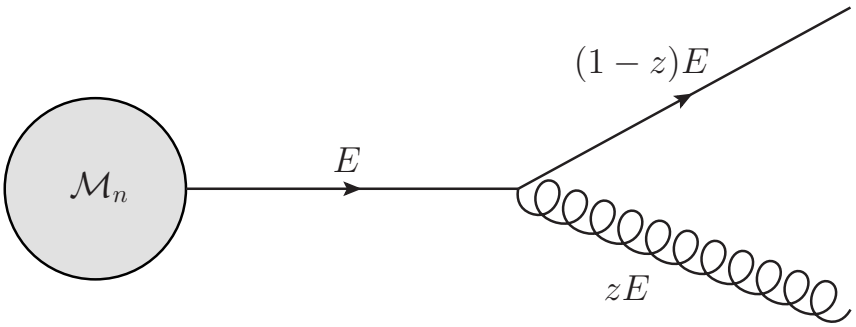



Figure 1.2: Kinematics for the splitting of one quark into a quark and gluon.

hadron beam picks up a flux factor, called a *Parton Distribution Functions* (PDF), which gives the probability of finding one specific parton with some fraction $x$ of the hadronic energy at some factorisation scale $\mu_f$. PDFs are universal, process-independent, thus they can be fitted from data directly. The evolution from one scale to another is possible in perturbative QCD via the DGLAP equations [19–21]. For the rest of the thesis we will focus on perturbative calculations, and ignore the finer details of PDFs. For numerical simulations we will take them as an input from the standard library LHAPDF [61].

Both $\mu_r$ and $\mu_f$ are not physical quantities, but merely "mathematical tricks"; all order calculations are independent of them. At fixed order on the other hand, changing the scale does give a subleading effect. Consequently higher-order calculations are generally more stable under scale variation. We can exploit such variation to get an estimate of the neglected subleading effects. Typically this is done by varying $\mu_r$ and $\mu_f$ by factors between 0.5 and 2. This is only a convention, and one could also decide on a larger or smaller range. Furthermore scale variations should not be confused with probabilistic uncertainties; they only illustrate the perturbative stability of calculations. One should therefore be careful when selection a scale, to not artificially tune predictions to observables. We will discuss scale choice again in the later chapters when showing results from actual simulations.

### 1.4.2 IR divergences

In the previous section we only considered UV divergences. At the infrared end *soft* and *collinear* divergences appear, which we will describe in the following. As an example we consider the quark propagator for the $q \to qg$ branching from fig. 1.2 in the massless limit

$$\frac{\not{p}_q + \not{p}_g}{(p_q^\mu + p_g^\mu)^2} = \frac{\not{p}_q + \not{p}_g}{2E_q E_g(1-\cos\theta)} \tag{1.4.10}$$

$$= \frac{\not{p}_q + \not{p}_g}{2E^2 z(1-z)(1-\cos\theta)}\,, \tag{1.4.11}$$

where $\theta$ is the angle between the outgoing quark and gluon, and $z = {}^{E_g}/_{E}$ is the energy fraction of the gluon. When either the gluon ($z \to 0$) or the quark ($z \to 1$) is soft, or both are collinear ($\theta \to 0$) the propagator diverges.

Physically we can not observe arbitrarily soft partons, since every detector will have a finite energy threshold. Neither will we be able to resolve two partons with arbitrarily small angles, eventually two close particles will be indistinguishable from one particle with the combined momenta. To circumvent such miss interpretations, one has to use a *IR safe* definition of measured particles, called *jets*, which we will define in the next section. In such observables we integrate over the unresolved, soft region. The resulting phase space has the same multiplicity as the virtual correction with one extra loop, and contributes at the same order in $\alpha_s$. It was proven in the KLN theorem [62–64] that both the real and virtual terms in the matrix element have the same pole structure for any order in $\alpha_s$, but with opposite signs. The resulting cross section only picks up finite corrections, without divergences.

In numerical implementations these divergences are challenging since the poles lie in different phase spaces. It is therefore common to introduce a counter-term to cancel the divergences. For example at NLO we expand

$$\sigma^{\mathrm{NLO}} = \int \mathrm{d}\Phi_N \left[\mathcal{B}_n + \mathcal{V}_n + \mathcal{I}_n + \int \mathrm{d}\Phi_1 \left(\mathcal{R}_{n+1} - \mathcal{C}_{n+1}\right)\right], \tag{1.4.12}$$

where $\mathcal{B}_n$ is the LO contribution. The subtraction term $\mathcal{C}_{n+1}$ has to be chosen, such that its difference to the real emission $\mathcal{R}_{n+1} - \mathcal{C}_{n+1}$ is finite, while still being simple enough to be integrated analytically to calculate $I_n$ from

$$\int \mathrm{d}\Phi_N \left(\mathcal{I}_n - \int \mathrm{d}\Phi_1 \mathcal{C}_{n+1}\right) = 0. \tag{1.4.13}$$

Since the pole structure of the virtual correction and the real emission are identical, this implies that the sum of the virtual and integrated subtraction terms $\mathcal{V}_n + \mathcal{I}_n$ is also finite. Equation (1.4.12) is generic and there are multiple subtraction schemes with different choices for $\mathcal{C}_{n+1}$ [45–47].

In general, even after subtraction, NLO (or higher) calculations pick up negative weights in Monte Carlo integrations (see section 2.4). Consequently, they require higher statistics, on top of an already more complicated integrand, which increases the considerable computational cost [65, 66].

### 1.4.3 Parton showers

Instead of treating the IR poles order by order in the strong coupling, one can also factorise the matrix element and treat the IR limit at all orders. To illustrate this we again consider the $q \to qg$ splitting from fig. 1.2. When the outgoing gluon goes soft both quarks go on-shell. We can therefore repeat the branching, by attaching another (soft) gluon to the outgoing quark. Each of these splittings only depends on the previous state

$$\mathrm{d}\sigma_{n+1} \to \mathrm{d}\sigma_n \frac{\mathrm{d}t}{t} \mathrm{d}z \frac{\alpha_s}{2\pi} P_{q\to qg}(z) \equiv \mathrm{d}\sigma_n \mathrm{d}t\, \mathrm{d}z \mathcal{P}_{q\to qg}(t,z), \tag{1.4.14}$$

where $P_{q\to qg}$ is the splitting kernels as appearing in the DGLAP equations, and $t$ is the evolution parameter [67]. For the other QCD splittings, $g \to gg$ and $g \to q\bar{q}$, one would find the same expression with different kernels $P_{a\to bc}$. The exact definition of $t$ can vary between different implementations, as long as it evolves the shower from some high scale down to the IR limit [26]. For example Pythia [28] and Sherpa [29] use (different) $p_\perp$ based definitions, while Herwig [27] uses angular ordering.

Equation (1.4.14) gives the probability of one emission. The virtual correction follows from unitarity, i.e. the chance of having no splitting in the infinitesimal region $\mathrm{d}t$ is $1 - \mathrm{d}t\mathcal{P}$. Equivalently, the chance of having the first splitting at scale $t$ is

$$\frac{\mathrm{d}\Delta(t_0,t)}{\mathrm{d}t} = \Delta(t_0,t) \int_{z_{\min}}^{z_{\max}} \mathrm{d}z \mathcal{P}(t,z), \tag{1.4.15}$$

where $\Delta(t_0,t)$ is the probability of having no splitting between scales $t_0$ and $t$. In general the soft regulation cuts $z_{\min}$ and $z_{\max}$ can depend on $t$ and $t_0$. Solving eq. (1.4.15) implies that the no-splitting probabilities are *Sudakov form factors*

$$\Delta(t_0,t_1) = \exp\left(-\int_{t_1}^{t_0} \mathrm{d}t \int_{z_{\min}}^{z_{\max}} \mathcal{P}(t,z)\right). \tag{1.4.16}$$

When multiple splittings are possible, e.g. $g \to gq$ and $g \to q\bar{q}$, they are added for the branching or multiplied for the no-branching probabilities respectively.

To numerically solve eq. (1.4.15) one uses parton showers. A parton shower produces different splittings at scales $t_i$ with $t_i \leq t_{i+1}$. The initial scale $t_0$ follows from a fixed order matrix element, while the final scale $t_N$ is the resolution parameter. Smaller $t_N$ explore softer regions, leading to more emissions. For sufficiently small values of $t_N$, IR safe observables must not be sensitive to it. Again we cannot measure only one parton, but we have to consider all partons inside the soft and collinear cascade. One of these cascades has to be *clustered* together to form one *jet*.

Reference 68 provides an excellent review of different jet definitions, all of which are implemented in the standard library `FastJet` [69]. At the LHC the anti-$k_t$ algorithm [70] is almost exclusively used, since it produces—experimentally desired—rounder, more regular shaped jets compared to other sequential clustering algorithms. The idea is to undo parton shower by repeatedly combining two particles $i$ and $j$ together in one jet, to get closer to the hard process. At each step, particle $i$ and $j$ are chosen such that they have the minimal distance

$$d_{ij} = \min\left(p_{i\perp}^{-2}, p_{j\perp}^{-2}\right) \frac{\Delta R_{ji}^2}{R^2} \tag{1.4.17}$$

of all possible pairs, with the jet radius

$$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2, \tag{1.4.18}$$

azimuthal angle $\phi_i$ and rapidity

$$y_i = \frac{1}{2} \ln\left(\frac{E_i + p_i^z}{E_i - p_i^z}\right). \tag{1.4.19}$$

If any (clustered) jet falls below the beam distance $d_{iB} = p_{i\perp}^{-2}$ it is removed from the list of jets before the next iteration. The jet radius $R$ acts as an collinear cut; with small $R$ more and softer jets are produced.

Compared to fixed order calculations, parton showers are not limited by the number of emissions. Instead eq. (1.4.16) includes infinitely many powers of $\alpha_s$, at the cost of assuming every emission to be in the IR limit. As such parton showers *resum* the leading soft and collinear logarithms. Current shower implementations are at least leading-logarithm accurate with full NLL in active development [71–75]. Parton showers are also key in Monte Carlo simulation to interpolate between hard and non-perturbative physics. As such they take the inclusive cross section and the general kinematics from fixed-order calculations—for example by matching with NLO [76,77]—and distribute the momenta over many partons. At the lower cut $t_N = t_{\text{cut}}$, *hadronisation* takes over to generate and decay hadrons, which are observable particles.

# Chapter 2

# Resummation in QCD

In the previous section we discussed how general kinematics are well described by fixed-order calculations. We further saw how soft and collinear logarithms require a resummation from parton showers, which originated from the DGLAP equations. Similarly to capture the dependence on the cross-section the jet radius $R$ a resummation of $\alpha_s^n \ln^n 1/R$ is required [78, 79]. Both examples act in the IR limit, though resummation itself is more general and can be required in various regions of phase space. In the following chapter we will introduce the high energy limit, where logarithms $\log|\hat{s}/\hat{t}|$ become large. They can be resummed, using e.g. the BFKL equation [22–25]. We will focus on the formalism of High Energy Jets. HEJ is a Monte Carlo generator, which also includes the BFKL logarithms, but unlike BFKL the expansion is done on a matrix element level, not the phase-space.

## 2.1 High energy limit

When we consider the high energy limit, we formally take the *multi-Regge kinematic* (MRK) limit [80]

$$y_i \ll y_{i+1} \quad \wedge \quad |p_{i\perp}| \sim p_\perp \qquad \forall i \tag{2.1.1}$$

$$\Leftrightarrow \quad \hat{s} \gg \hat{s}_{ij} \gg p_\perp \qquad \forall i,j \tag{2.1.2}$$

for $2 \to n$ QCD processes. Here and throughout the rest of this thesis we always label all particles according to their rapidity such that $y_i \leq y_{i+1}$. In the MRK limit the Mandelstam variables simplify to

$$\hat{s}_{ij} = -\hat{u}_{ij} = p_\perp^2 \exp\left|\Delta y_{ij}\right| \tag{2.1.3}$$

$$\hat{t}_{ij} = -p_\perp^2. \tag{2.1.4}$$

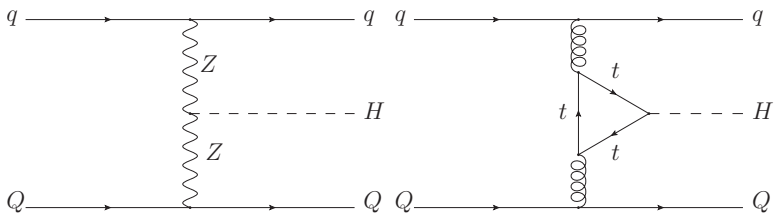



Figure 2.1: Representative LO Feynman diagrams for the production of $qQ \to qHQ$ through weak-boson (left) and gluon fusion (right).

Thus requiring high partonic energies is equivalent to requiring large rapidity separations between every final state parton. Similarly to the soft/collinear limit we pick up large logarithms $\log |\hat{s}/\hat{t}| \to \log \hat{s}/p_\perp^2 \sim \Delta y$. Therefore any observable which depends on a large rapidity separations between jets is sensitive to high energy logarithms.

An example of an observable where these logarithmic sensitivities occur is the Higgs boson plus dijet process. Typically one wants to distinguish the two production channels, electroweak production from *weak-boson fusion* (also called *vector*-boson fusion, VBF) and the QCD production from *gluon fusion* (GF). In gluon fusion the Higgs boson couples to a heavy quark (i.e. top) loop to the gluon, thus at Born-level it contributes to order $\alpha_s^4$. On the other hand VBF comes from a tree level Feynman diagram, but is suppressed with the weak coupling $\alpha_w^2$. Even though both channels can produce the same final state with only quarks (see fig. 2.1), quantum interference between them is negligible [81–83]. It is therefore reasonable to study the two processes independently.

Naively one might expect both processes to be equally likely, since $\alpha_s^2 \sim \alpha_w$. But in reality at the LHC gluon fusion is twice as likely as VBF, due to the proton PDF being dominated by low-energy gluons. Hence to measure the electroweak production, one has cut away the low energy region, commonly by a cut on the invariant mass of the hardest two jets [85–87]

$$m_{12}^2 = \hat{s}_{12} = 2p_{1\perp}p_{2\perp}\left[\cosh\left(y_1 - y_2\right) - \cos\left(\phi_1 - \phi_2\right)\right], \tag{2.1.5}$$

where we ignored the masses of the jets, i.e. eq. (2.1.5) is the LO approximation. The cross section from gluon fusion is highly peaked at small $m_{12}$ comparably to the (almost) flat distribution from VBF (see fig. 2.2). After requiring $m_{12} > 400\,\text{GeV}$ and $|y_1 - y_2| > 2.8$ their relative contributions flip; VBF becomes at least twice as likely.

These cuts are used more generally even for multijet production together with a $Z$ [88] and $W$ boson [89]. Yet $m_{12}$ is affected by multiple logarithms—soft/collinear, coming from $p_\perp$, and high energy, from $\Delta y$—which makes it hard to accurately

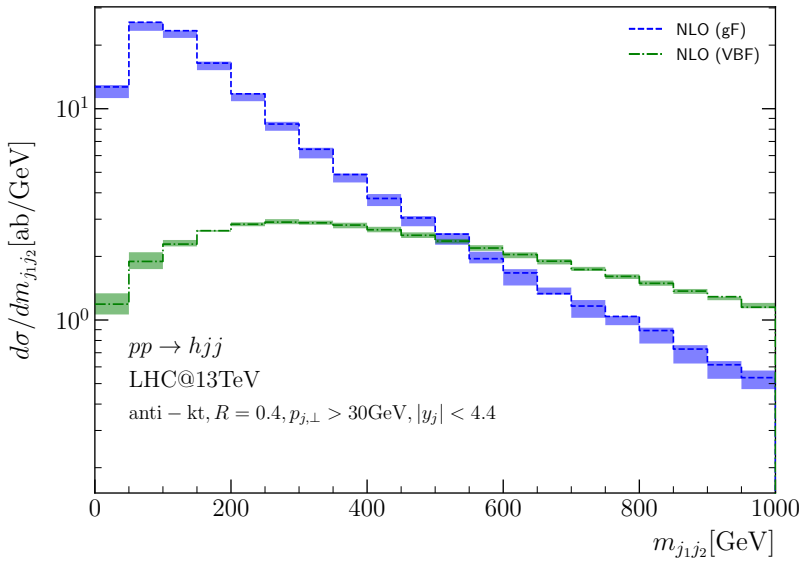



Figure 2.2: Invariant mass between the hardest jets at NLO for QCD (blue, dashed) and electroweak (green, dash-dotted) Higgs boson plus dijet production. The NLO results where generated with MCFM [54, 84]. From ref. 85.

predict the differential cross section $\mathrm{d}\sigma/\mathrm{d}m_{12}$. In ref. 88 modelling $m_{12}$ only with parton showers lead to an 11% systematic uncertainty (out of a total 17%) for the measurement of electroweak $Z$-boson production. Similarly, for the production of a single $W$ boson alongside multiple jets at 7 TeV, ATLAS observed a large spread between the predictions of different parton showers at large rapidity separations [90]. Controlling the high energy logarithms, via a resummation in the MRK limit, can stabilise the predictions and reduce the theoretical uncertainty.

More direct measurements of the perturbative corrections in the high energy limit were done by the D0 Collaboration for $W$-boson plus dijet production at a centre of mass energy of 1.96 TeV [91]. They measured the probability of emitting an additional jet depending on the rapidity separations between different jet pairs (see left of fig. 2.3). The emission probability grows as the maximal rapidity span of the jets, $\Delta y_{fb}$, increases. Partially this is explained by opening up the phase space, since there is little room for the third jet at $\Delta y_{fb} = 0$. Therefore up to rapidity separations of 3.5 the parton shower result from SHERPA [92] increases. Afterwards the phase space is large enough to include all collinear emissions, thus SHERPA predicts no further increase. The measured probability on the other hand keeps on rising linearly up to the maximal range $\Delta y_{fb} = 6$. This clearly shows an enhancement of radiation with $\Delta y_{fb}$.

Conversely if we look at the same observable, but for the two hardest jets and

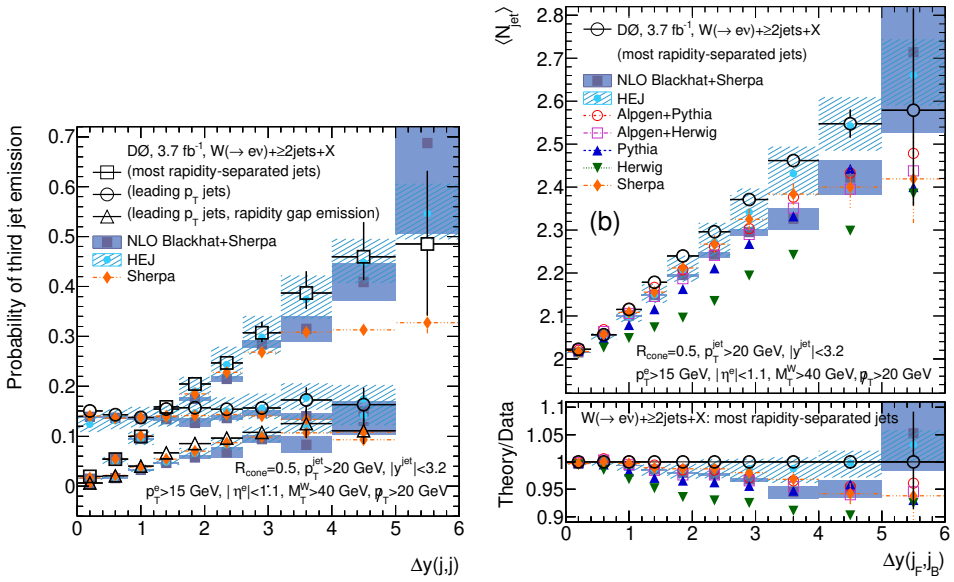



Figure 2.3: Measurement of the third jet emission rates for $W$ with two jets at the D0 detector. Left: Emission probability dependent on the rapidity separation between the forward and backwards (boxes), and the hardest jets (circles). Additionally the emission of a third jet inside the hardest emission is shown (triangles). Right: Average number of jets dependent on the rapidity between the extremal jets. From ref. 91

allow the third jet to be anywhere in phase space, the probability distribution is flat (circles). Extra emissions are not enhanced, but evenly distributed as expected for soft or collinear emissions. If we only look at the jets in the rapidity gap the linear increase is visible again (triangles), though the total separation is smaller and high energy effects less prominent.

The NLO calculations from `BlackHat`+SHERPA [48] do catch the rise of the probability with $\Delta y$ better then SHERPA alone. However, when compare the average number of jets against $\Delta y_{fb}$ (right side of fig. 2.3) NLO does not reach the required multiplicity, and thus, like parton showers, underestimates the number of jets. The linear growth of the number of jets with $\Delta y$ is characteristic for BFKL evolution [93]. Similar measurements were performed by the ATLAS collaboration for pure jets [94, 95]. Only when the high energy BFKL logs are included in predictions, as is the case for the HEJ program, is the data described well over the full range of $\Delta y$.

## 2.2 Regge theory

These experimental observations can be explained theoretically through *Regge theory*. In the following we give a heuristic overview, a more detailed introduction can for

example be found in ref. 96. Historically Regge theory was used to describe hadron–hadron interactions, even predating QCD. It predicts that a scattering amplitude scales as [80]

$$\mathcal{M} \sim \gamma \prod_i s_{i,i+1}^{\alpha_i}, \tag{2.2.1}$$

where $\alpha_i$ is the spin of the *Reggeons* $i$, and $\gamma$ some residue depending only on transverse momenta. Reggeons are effectively $t$-channel exchange particles. Each additional emission $i$ gives an enhancement proportional to its energy transfer $s_{i,i+1}$ for $y_i \leq y_{i+1}$.

Most remarkably, eq. (2.2.1) can also be derived in QCD [97], where instead of a Reggeon we have $t$-channel quark or gluon exchange. Explicitly for a Higgs-boson production with two jets at rapidities $y_1$ and $y_2$, and a rapidity hierarchy between the jets and the Higgs boson ($y_1 \ll y_H \ll y_2$), we get [98]

$$\mathcal{M} \sim s_{1H}^{\alpha_1} s_{H2}^{\alpha_2} \gamma \left( t_1, t_2, \frac{s_{12}}{s_{1H} s_{H2}} \right), \tag{2.2.2}$$

where $t_i$ is the momentum squared flowing through the $t$-channel after particle $i$. The dominant contribution is the exchange of one $t$-channel gluon (spin-1 exchange), as the exchange of a quark would be suppressed by $\sqrt{s_{qH}}$. In particular the general pattern of eqs. (2.2.1) and (2.2.2) holds for arbitrarily high multiplicities, giving a relation between a rapidity ordering and the scaling of an scattering amplitude.

As an example, we consider the production of two quarks, one Higgs boson and one gluon (see fig. 2.4). In the dominant configuration both quarks are the most forward and backwards particles (rapidity ordering $qQ \to qgHQ$), which we can represent as a non-crossing, planar $t$-channel gluon exchange where all partons are sorted according to their rapidity. This matrix element scales like $|\mathcal{M}_1|^2 \sim s_{12}^2 s_{2H}^2 s_{H3}^2 \gamma_1$, where as before $\gamma_1$ only depends on the transverse momenta. Hence $|\mathcal{M}_1|^2 / s^2$ becomes constant in the high energy region (red line on the left of fig. 2.4). We call this the *Fadin–Kuraev–Lipatov* configuration. At leading logarithmic accuracy there is no other rapidity ordering.

If we change one of the quarks with the gluon ($\mathcal{M}_2$, blue line in fig. 2.4) we get an unordered gluon emission. A planar pure $t$-channel diagram now requires a quark (spin-$\frac{1}{2}$) between the gluon and the backward quark. The amplitude squared $|\mathcal{M}_2|^2 \sim s_{12} s_{2H}^2 s_{H3}^2 \gamma_2$ is suppressed by one power of $s_{12}$ compared to $|\mathcal{M}_1|^2$, or equivalently exponentially suppressed in $\Delta y_{12}$. This is the first subleading contribution. We can explicitly multiply with $s_{12}$ to reconstruct the flattening of $\mathcal{M}_2$ (see right side of fig. 2.4).

At last we also swap the Higgs boson with the first quark jets ($qQ \to gHqQ$). The

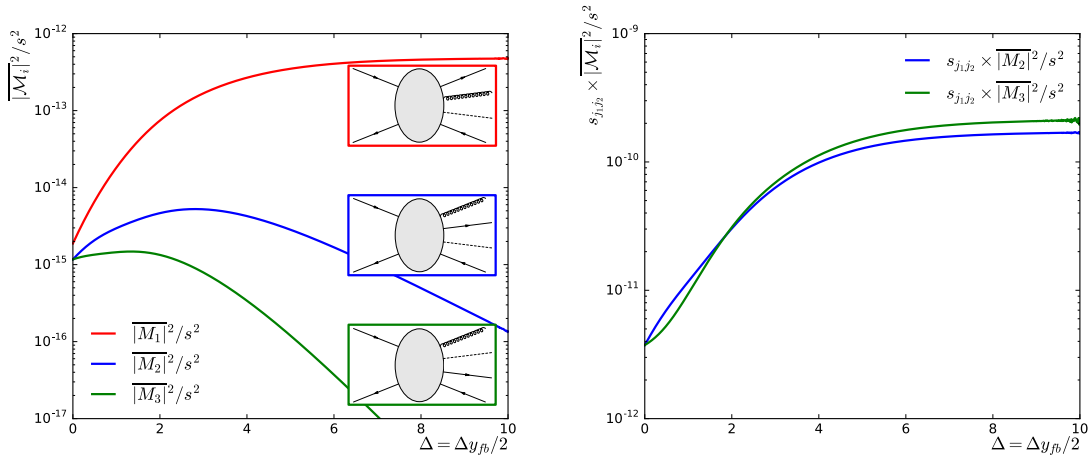



Figure 2.4: High energy limit of leading order matrix element for different rapidity ordering of the $qgHQ$ final state. Both quarks $q$ and $Q$ have different flavour to uniquely identify them. The Feynman diagrams shown on the left correspond to the rapidity ordering, i.e. $qQ \to qgHQ$ (red), $qQ \to gqHQ$ (blue) and $qQ \to gHqQ$ (green). On the right the subleading matrix elements $qQ \to gqHQ$ and $qQ \to gHqQ$ are multiplied by the invariant mass between the gluon and the first quark. From ref. 98

matrix element then scales like $|\mathcal{M}_3|^2 \sim s_{1H} s_{H2} s_{23}^2 \gamma_3$ (green in fig. 2.4). This ordering is now suppressed by both $s_{1H}$ and $s_{H2}$. According to eq. (2.1.3) this is equivalent to a suppression with $s_{12}$. Consequently $\mathcal{M}_3$ also contributes at next to leading logarithmic order.[1] In fact we could have ignored the Higgs boson in the discussion above, by only considering the colour connection between final state partons. As the Higgs boson is colour-neutral it does not affect the colour exchange. We always either have a colour-octet or colour-singlet, corresponding to the gluon or quark exchange respectively.

At leading colour every colour-flow that can be represented by *two-sided plots* contributes to the MRK limit, as long as the colour and rapidity order coincide [100]. In the first row of fig. 2.5 we illustrate multiple allowed leading-colour connections for three FKL processes. Incoming partons are shown at the top or bottom, and outgoing on the left or right. The outside arrows represent the (anti-)colour flow. Each gluon can couple to any of the two sides. In particular figs. 2.5a and 2.5b represent two (equally likely) colour flows for $gg \to ggg$, where gluon 2 is connected either to the colour or anti-colour of the exchanged gluon. Such crossings are allowed

[1] This actually depends on the definition of the resummation. We assume a large rapidity gaps *between partons* $y_1 \ll y_2$. Hence here we only count $s_{12}$. One could take $s_{1H}$ and $s_{2H}$ separately when requiring $y_1 \ll y_H \ll y_2$, which would make $\mathcal{M}_3$ contribute to N$^2$LL. Obviously the first approach is more general.

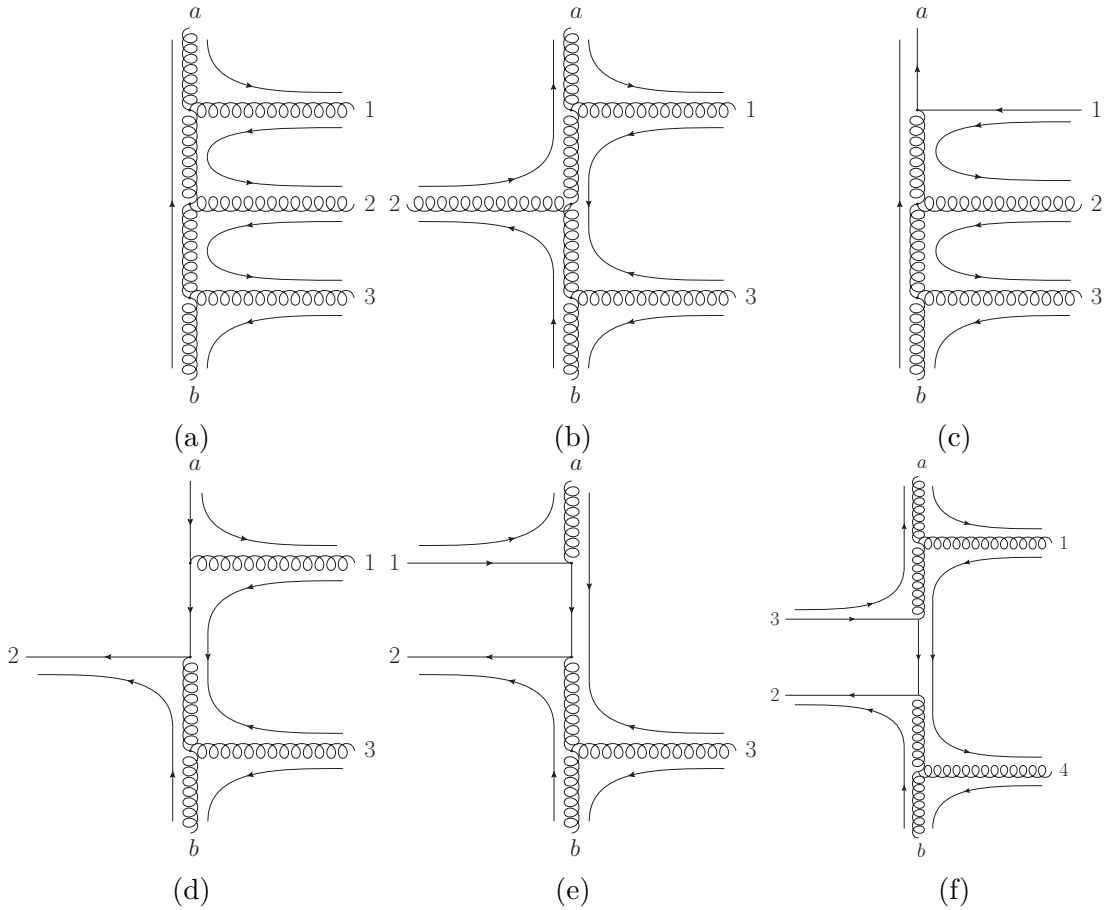



Figure 2.5: Different colour configurations for multi-jet processes in the MRK limit. Incoming partons $a$ and $b$ are shown on the top and bottom, and outgoing partons on the left and right, with their indices sorted according to rapidity, $y_i \leq y_{i+1}$. The outside arrows represent the colour flow. All processes in the first row contribute at leading logarithm, wile the second row shows some NLL processes at lowest multiplicity. The subleading processes are the emission of an unordered gluon further backwards then the first quark (d), one $\bar{q}q$ pair most backwards in rapidity (e), and an central $q\bar{q}$ pair inside the FKL chain (f). (a) to (e) are $t$-channel and (f) is a $u$-channel exchanges respectively, i.e. in (f) the colour exchange is flipped compared to the rapidity at the $q\bar{q}$ pair.

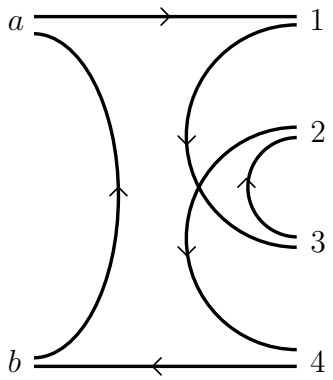



Figure 2.6: Illustration of a subleading-colour configuration. The crossing between 2 and 3 can only be resolved by breaking the rapidity order. From ref. 99.

as long as they do not break the rapidity ordering. In contrast crossings like fig. 2.6 are forbidden, since the flow can only be untangled by swapping 2 and 3 vertically, which violates the rapidity order.

Any process that cannot be represented by only colour-octet exchanges is subleading. At NLL up to one colour-singlet exchange is allowed (see second row of fig. 2.5). We already saw the unordered emission in fig. 2.4, which is now shown in fig. 2.5d. However there are two more diagrams; containing a $q\bar{q}$ pair either extremal or central in rapidity, corresponding to figs. 2.5e and 2.5f respectively. Extremal $q\bar{q}$ pair are related to unordered emissions by exchanging the incoming quark with the unordered gluon. Central $q\bar{q}$ pair on the other hand is a radiative correction to the extremal $q\bar{q}$ pair, only contributing at four or more partons. Subleading contributions correspond to the quasi-MRK limit [101]

$$y_1 \ll \cdots \ll y_{i-1} \ll y_i,\ y_{i+1} \ll y_{i+2} \ll \cdots \ll y_n\,. \tag{2.2.3}$$

Compared to the strict MRK limit eq. (2.1.2), two partons $i$ and $i+1$ can be close in rapidity. The $q\bar{q}$ pairs acts as a single "blob", which couples to the $t$-channel gluon. Consequently the colour ordering within that blob does not matter. Both the $t$- and $u$-channel exchange in the $q\bar{q}$ pair shown on the figs. 2.5e and 2.5f respectively contribute. There is one more subleading configuration where two gluons are close to each other. We do not show this processes here, the possible colour configuration are similar to the FKL configurations up to an additional $u$-channel exchange.

The identification between $t$-channel exchange and colour structure connects MRK to the Parke-Taylor formula [102]. To leading order in the high energy limit each

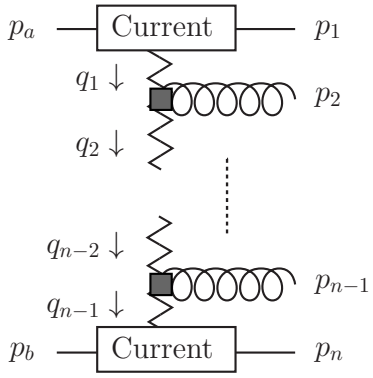



Figure 2.7: Schematic representation of the HEJ amplitude eq. (2.3.5). The boxes are the impact factors or *currents*, which are connected through a reggeised gluon (zigzag line). Each emission in the rapidity gap comes from a Lipatov vertex (grey boxes). From ref. 98.

additional gluon emission then picks up a factor [98, 103, 104]

$$\overline{|\mathcal{M}_{n+1}|}^2 \to \frac{16\pi\alpha_s C_A}{p_{i\perp}^2} \cdot \overline{|\mathcal{M}_n|}^2. \tag{2.2.4}$$

Through this factorisation we can build up the matrix element recursively. Any $n+1$ emission is a radiative correction to the $n$-parton base-process. Integrating over the extra emission phase space leads to

$$\int_{y_{i-1}}^{y_{i+1}} \frac{\mathrm{d}y_i}{4\pi} \int \frac{\mathrm{d}^2 p_\perp}{(2\pi)^2} \frac{16\pi\alpha_s C_A}{p_\perp^2} = 4\alpha_s C_A \Delta y_{i-1,i+1} \int \frac{\mathrm{d}^2 p_\perp}{(2\pi)^2} \frac{1}{p_\perp^2}\,. \tag{2.2.5}$$

The last integral diverges in the soft limit $p_\perp \to 0$, which we already saw in section 1.4.2. In section 2.3.3 we will explicitly show how this IR pole cancels when we include virtual corrections. Important here is that the matrix element became independent of the rapidity, which leads to an enhancement of the cross section with $\Delta y \sim \log|\hat{s}/\hat{t}|$.

## 2.3 High Energy Jets

In the previous section we saw the high energy logarithms appearing in QCD. We concluded with a factorisation of the matrix element. Higher multiplicity matrix elements can therefore be constructed from "simple" building blocks. In this section we will introduce the framework of High Energy Jets to resum logarithms of the form $\log|\hat{s}/\hat{t}|$. In HEJ the resummation is done by explicitly integrating over the phase space with the exclusive matrix element for each multiplicity, and adding all of these terms together. All approximations are applied to the matrix element itself;

the phase space has to be integrated numerically. We will only describe the HEJ matrix element here, while postponing the phase space integration to chapter 3.

We start with a generic $2 \to 2$ (pure-QCD) matrix element in the MRK limit [98]

$$\left|\mathcal{M}_{f_a(p_a)f_b(p_a)\to f_1(p_1)f_2(p_2)}\right|^2 \to \hat{s}^2 \frac{|C(p_a,p_1)|}{(p_a-p_1)^2} \frac{|C(p_b,p_2)|}{(p_b-p_2)^2}\,, \tag{2.3.1}$$

where the $f_i$ represent the flavour of the external particle at leg $i$. As before all outgoing particles are numbered according to their rapidity ($y_i < y_{i+1}$); $a$ and $b$ are the incoming partons with negative and positive $p_z$ respectively. Each impact factor $C(p_i, p_j)$ only depends on the momentum of one leg in a $t$-channel diagram. At leading logarithm we do not have a particle change at the impact factors, $f_a = f_1$ and $f_b = f_n$.

Using factorisation, we can write matrix element for $2 \to 3$ processes as

$$\begin{aligned}\left|\mathcal{M}_{f_af_b\to f_1f_2f_3}\right|^2 \to\ & \hat{s}^2 \frac{|C(p_a,p_1)|}{(p_a-p_1)^2} \frac{|V_L(p_2)|^2}{(p_a-p_1)^2\,(p_b-p_3)^2} \frac{|C(p_b,p_3)|}{(p_b-p_3)^2} && (2.3.2)\\ =\ & \hat{s}^2 \frac{|C(p_a,p_1)|}{t_1} \frac{|V_L(p_2)|^2}{t_1t_2} \frac{|C(p_b,p_3)|}{t_2}\,, && (2.3.3)\end{aligned}$$

where the emission of an extra parton is done through the *Lipatov vertex* $V_L$ [105,106]. In the MRK limit the second parton has to be a gluon ($f_2 = g$). Hence the Lipatov vertex is an effective term for the emission of one gluon from a *reggeised*, $t$-channel gluon. According to eq. (2.2.4) in the high energy limit

$$\frac{|V_L(p_2)|^2}{t_1t_2} \to \frac{4}{p_{2\perp}^2}\,. \tag{2.3.4}$$

We can generalise eq. (2.3.3) for arbitrary many gluons, which gives the ansatz for the generic, colour and spin averaged matrix element for HEJ at tree-level

$$\begin{aligned}\overline{\left|\mathcal{M}^{\text{HEJ,tree}}_{f_af_b\to f_1g\cdots gf_n}\right|^2} =& \frac{1}{4\,(N_C^2-1)}\ \left\|\mathcal{S}_{f_af_b\to f_1f_n}\right\|^2 \left(g_s^2\ K_{f_1}\ \frac{1}{t_1}\right)\left(g_s^2\ K_{f_n}\ \frac{1}{t_{n-1}}\right)\\ &\cdot \prod_{i=1}^{n-2}\left(\frac{-g_s^2C_A}{t_it_{i+1}}\ V_L^\mu(q_i,q_{i+1})V_{L\mu}(q_i,q_{i+1})\right),\end{aligned} \tag{2.3.5}$$

where the impact factors are grouped together in $\mathcal{S}_{f_af_b\to f_1f_n}$, the double bars indicate a sum over all spins. $K_{f_i}$ are the appropriate colour factors for the coupling of the $t$-channel gluon to the impact factors. We use the momenta flowing in and out along the gluon exchange $q_{i-1}$ and $q_i$ as arguments of the Lipatov vertices (see fig. 2.7)

$$q_1 = p_a - p_1, \quad q_{i+1} = q_i - p_{i+1}\,. \tag{2.3.6}$$

This choice does not affect the limit in eq. (2.3.4) since $q_i^2 = t_i$.

### 2.3.1 Simplest HEJ current

$\mathcal{S}_{f_a f_b \to f_1 f_n}$ is the only process-specific term of eq. (2.3.5); it represents the $2 \to 2$ amplitude. For example the resummation for $gg \to gg$ and $qQ \to qQ$ only differ in $\mathcal{S}_{gg\to gg}$ and $\mathcal{S}_{qQ\to qQ}$. Furthermore we already saw in eq. (2.3.1) that the matrix elements factorise into impact factors, or *currents*.

The simplest process in HEJ is the scattering of two different-favoured quarks $qQ \to qQ$, where only one $t$-channel diagram contributes. This amplitude is built from two terms (see eqs. (A.1.18) and (A.1.20))

$$\overline{\mathcal{M}_{qQ\to\, qQ}} = \frac{1}{4t} \sum_{h_1,h_2\in\{+,-\}} j_\mu^{h_1}(p_a,p_1) j^{h_2\,\mu}(p_b,p_n), \tag{2.3.7}$$

where we defined the two currents

$$j_\mu^-(p,q) \equiv \langle p|\mu|q] \equiv j_\mu^+(q,p). \tag{2.3.8}$$

Comparing eq. (2.3.7) with eq. (2.3.5) we can read off the (contracted) HEJ currents for $qQ \to qQ$

$$\left\|\mathcal{S}_{qQ\to\, qQ}\right\|^2 \equiv 4t_1 t_{n-1} \overline{\left|\mathcal{M}_{qQ\to\, qQ}\right|}^2 \tag{2.3.9}$$

$$= \sum_{h_1,h_2\in\{+,-\}} \left[j_\mu^{h_1}(p_a,p_1) j^{h_2\,\mu}(p_b,p_n)\right]^2. \tag{2.3.10}$$

We wrote $t_1$ and $t_{n-1}$ in expectation of generalising this result to arbitrary multiplicities, obviously for $2 \to 2$ processes $t_1 = t_{n-1} = t$. The same calculation also holds for anti-quark currents. The colour factors are

$$K_q = K_{\bar{q}} = C_F\,. \tag{2.3.11}$$

Each current only depends on the two momenta at one side of the $t$-channel diagram, without any $s$- or $u$-crossing[2]. In the MRK limit this is guaranteed by eq. (2.3.1). Consequently even for the scattering of two same-flavoured quarks we can use eq. (2.3.10), i.e. $\mathcal{S}_{qq\to qq} = \mathcal{S}_{qQ\to qQ}$. This approximation only includes some parts of the full LO result, which is allowed as long as we catch the right LL behaviour. In section 3.1 we will detail the method of *matching* HEJ with fixed-order calculations to get both LO and LL correct results.

In the high energy limit the only difference between external gluons and quarks

[2]In this context $t$- and $u$-channel "diagrams" refers to the colour flow from the previous section. They should not be confused with Feynman diagrams. To calculate matrix elements we always include all possible Feynman diagrams; statements about individual Feynman diagrams are, in general, not gauge invariant.

is their colour factor $K_g = C_A$ [32]. We can use the same equations as before for the gluon currents. As shown in ref. 33 for helicity conserving amplitudes, like $q^-g^+ \to q^-g^+$, we recover the full matrix element by using the *Colour Acceleration Multiplier* (CAM)

$$K_g\left(p_2^-, p_b^-\right) = \frac{1}{2}\left(C_A - \frac{1}{C_A}\right)\left(\frac{p_b^-}{p_2^-} + \frac{p_2^-}{p_b^-}\right) + \frac{1}{C_A} \to C_A\,, \tag{2.3.12}$$

where $p^- = E - p_z$ is the negative light-cone momentum with $p_b^- \to p_2^-$ in the MRK limit. Amplitudes with a helicity flip, like $q^-g^- \to q^-g^+$, have an extra term with a $u$- or $s$-channel pole. These poles are suppressed at high energies, thus we neglect them and instead use eq. (2.3.12) for all gluon currents. Explicitly, the tree-level HEJ matrix element for $gg \to gg$ is

$$\overline{\left|\mathcal{M}^{\text{HEJ,tree}}_{gg\to gg}\right|}^2 = \frac{1}{4\ (N_C^2-1)}\frac{g_s^4}{t^2}\ \left\|\mathcal{S}_{qQ\to qQ}\right\|^2 K_g(p_1^-, p_a^-)K_g(p_2^-, p_b^-), \tag{2.3.13}$$

which can be extended to arbitrary multiplicity through eq. (2.3.5).

So far we only considered $2 \to 2$ processes. We can expand our list of currents to also include the production of (one) weak boson. Flavour changing $W$ bosons can only couple to one of the quarks, which means we only need to replace one current

$$\mathcal{S}_{ud\to d\nu_l\bar{l}d} = j^\mu_W(p_a, p_1, p_\nu, p_{\bar{l}})j_\mu(p_b, p_2). \tag{2.3.14}$$

The full equation for $j^\mu_W$ can be found in ref. 107.

Since the $W$ boson changes the quark flavour it allows for an identification of the emitting leg. For $Z$-boson or photon production on the other hand both legs are identical, leading to an interference between them (and between $Z$ and $\gamma$) [108]

$$\mathcal{S}_{qQ\to ql\bar{l}Q} = j^\mu_{Z/\gamma}(p_a, p_1, p_l, p_{\bar{l}})j_\mu(p_b, p_2) + j_\mu(p_a, p_1)j^\mu_{Z/\gamma}(p_b, p_2, p_l, p_{\bar{l}}). \tag{2.3.15}$$

The dominant production channel of Higgs bosons is gluon fusion, where the Higgs boson couples to any gluons via a top-quark loop. In the *Higgs Effective Field Theory* (HEFT) one assumes infinite top-mass, such that the top-loop induced coupling between a Higgs and a gluon (or photon) reduces to a single vertex. Hence, in HEJ we get a current with the Higgs-boson coupling to the $t$-channel gluon [98]

$$\mathcal{S}_{qQ\to qHQ} = j_\mu(p_a, p_1)V_H^{\mu\nu}(q_1, q_2)j_\nu(p_b, p_2), \tag{2.3.16}$$

with the tensor

$$V_H^{\mu\nu}(q_i, q_j) = \frac{\alpha_s}{3\pi v}\left(g^{\mu\nu}q_i\cdot q_j - q_j^\mu q_i^\nu\right), \tag{2.3.17}$$

where $v$ is the vacuum expectation value. Equation (2.3.17) is the HEFT equivalent for triangle top-quark loops. The $1/m_t$ pole in the loop thereby cancels with the $m_t$ dependence in the Higgs coupling to give the effective coupling eq. (2.3.17) at $m_t \to \infty$. $V_H^{\mu\nu}$ is a correction to the reggeised gluon, whereas the $W$, $Z$ or photon cases above only changed the impact factors. For extremal gluons we get additional box diagrams, which lead to new impact factors when using the full top-quark loop [109]. We will present the amplitudes for Higgs-boson production with finite top-mass in chapter 4.

Since our ansatz is general enough to allow arbitrary $1 \to 2$ currents, we can exploit this further by including NLL configurations. As discussed in eq. (2.2.3) in the quasi-MRK limit we allow two partons to be close in rapidity. This leads to two new currents (unordered and extremal $q\bar{q}$), one vertex with a central $q\bar{q}$ pair (see fig. 2.5), and, in principle, both a vertex and an impact factor for two gluons close in rapidity. As first subleading corrections, the unordered emissions were calculated in ref. 98

$$\mathcal{S}_{qQ\to gqQ} = j^{\mu}_{\text{uno}}(p_a, p_1, p_2) j_{\mu}(p_b, p_3), \tag{2.3.18}$$

with $y_g < y_q$. The remaining two quark terms, central and extremal $q\bar{q}$ pair, for pure jets and $W$ bosons plus jets will be published soon [110]. This does not yet lead to full NLL accuracy in HEJ. But these currents are universal, such that we can combine them to include more complex configurations in the resummation[3], e.g.

$$\mathcal{S}_{qQ\to gqHQ} = j_{\text{uno }\mu}(p_a, p_1, p_2) V_H^{\mu\nu}(q_2, q_3) j_{\nu}(p_b, p_3). \tag{2.3.19}$$

In conclusion we only need a small set of currents to describe a wide range of processes.

### 2.3.2 Lipatov vertex

Following our general recipe in eq. (2.3.5) we now continue with the gluon emission vertices. These are key for reaching arbitrary multiplicities. Since we already saw that quarks and gluons are similar in the MRK limit, we only have to consider the process $qQ \to qgQ$ which is a radiative correction to $qQ \to qQ$. Each gluon can either be emitted from the external legs $\mathcal{M}_g$ or from the reggeised gluon $\mathcal{M}_g$ (see

[3]Not all subleading configurations can trivially be built this way. For example $qQ \to gHqQ$ requires a special $q \to gHq$ current $j^{\mu}_{\text{uno},H}$, which includes both an unordered gluon and a Higgs boson.

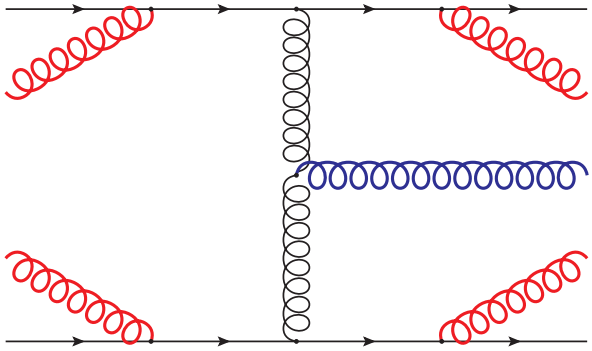


Figure 2.8: All five Feynman diagrams for the coupling of one gluon to $qQ \to qQ$ as included in the Lipatov vertex, each external gluon represents one possible diagram. The four diagrams in red are gluon emissions from the different quark legs, and the one in blue comes from the gluon-gluon vertex in the $t$-channel exchange.

fig. 2.8). The latter case is a pure $t$-channel exchange, as such it factorises to

$$\begin{aligned}\mathcal{M}_g =& f^{abc} T^c_{i_1 i_a} T^b_{i_3 i_b} \frac{-g_s^3}{t_1 t_2} j_\mu(p_a, p_1) j_\nu(p_b, p_3) \epsilon^*_\rho \\ & \cdot \left((q_1 + q_2)^\rho g^{\mu\nu} + (p_2 - q_2)^\mu g^{\nu\rho} - (q_1 + p_2)^\nu g^{\mu\rho}\right).\end{aligned} \tag{2.3.20}$$

In the high energy limit this simplifies further

$$\mathcal{M}_g \to f^{abc} T^c_{i_1 i_a} T^b_{i_3 i_b} \frac{g_s^3 \mathcal{S}_{qQ\to qQ}}{t_1 t_2} \epsilon^*_\mu \left(2p_a^\mu \frac{s_{23}}{\hat{s}} - 2p_b^\mu \frac{s_{12}}{\hat{s}} - (q_1 + q_2)^\mu\right). \tag{2.3.21}$$

For the emissions of the gluon from the external legs we assume the gluon to be soft, which gives

$$\begin{aligned}\mathcal{M}_q = -\mathrm{i} g_s^3 \mathcal{M}_{qQ\to qQ} \epsilon^*_\mu \Bigg( & T^a_{i_1 i} T^b_{i i_a} T^b_{i_3 i_b} \frac{p_1^\mu}{p_1 \cdot p_2} - T^b_{i_1 i} T^a_{i i_a} T^b_{i_3 i_b} \frac{p_a^\mu}{p_a \cdot p_2} \\ & + T^b_{i_1 i_a} T^a_{i_3 i} T^b_{i i_b} \frac{p_3^\mu}{p_3 \cdot p_2} - T^b_{i_1 i_a} T^b_{i_3 i} T^a_{i i_b} \frac{p_b^\mu}{p_b \cdot p_2} \Bigg),\end{aligned} \tag{2.3.22}$$

where we used the Eikonal approximation to the emission from the quark legs. Together with the MRK limit $p_a \to p_1$ and $p_b \to p_3$ we get

$$\mathcal{M}_q \to -f^{abc} T^c_{i_1 i_a} T^b_{i_3 i_b} \frac{g_s^3 \mathcal{S}_{qQ\to qQ}}{t_1 t_2} \epsilon^*_\mu \left(q_1^2 \frac{p_1^\mu}{p_1 \cdot p_2} - q_2^2 \frac{p_3^\mu}{p_2 \cdot p_3}\right). \tag{2.3.23}$$

Adding together eq. (2.3.23) and eq. (2.3.22) leads to the Lipatov vertex as used in the BFKL equation [106, 111]. To capture more of the kinematic behaviour, we do not use the identification $p_{a/b} \to p_{1/3}$; instead we apply eq. (2.3.22) directly. The

full amplitude is then

$$\mathcal{M}_{qQ\to qgQ} = f^{abc} T^c_{i_1 i_a} T^b_{i_3 i_b} \frac{g_s^3 \mathcal{S}_{qQ\to qQ}}{t_1 t_2} \epsilon^*_\mu V_L^\mu(q_1, q_2), \tag{2.3.24}$$

with the Lipatov vertex for a gluon emission at position $i+1$

$$\begin{aligned} V_L^\mu(q_i, q_{i+1}) = &- (q_i + q_{i+1})^\mu \\ &+ \frac{p_a^\mu}{2}\left(\frac{q_i^2}{p_{i+1}\cdot p_a} + \frac{p_{i+1}\cdot p_b}{p_a\cdot p_b} + \frac{p_{i+1}\cdot p_n}{p_a\cdot p_n}\right) + p_a \leftrightarrow p_1 \\ &- \frac{p_b^\mu}{2}\left(\frac{q_{i+1}^2}{p_{i+1}\cdot p_b} + \frac{p_{i+1}\cdot p_a}{p_b\cdot p_a} + \frac{p_{i+1}\cdot p_1}{p_b\cdot p_1}\right) - p_b \leftrightarrow p_n\,. \end{aligned} \tag{2.3.25}$$

As desired, eq. (2.3.25) is gauge invariant—which can explicitly be checked through the Ward identity, $p_{i+1}\cdot V_L = 0$—and fulfils the limit in eq. (2.3.4). With this we conclude the derivation of every building block we needed in eq. (2.3.5) to calculate HEJ matrix elements for arbitrary multiplicities at tree-level. So far we have only considered real emissions, which themselves contain IR singularities in soft limit. In the following section we will add virtual corrections to cancel these divergences.

### 2.3.3 Virtual corrections

In the MRK limit the virtual corrections can be included through the *Lipatov ansatz* [23]; each propagator in eq. (2.3.5) gets replaced with

$$\frac{1}{t_i} \to \frac{1}{t_i}\exp[\hat{\alpha}(q_{i\perp})(y_{i+1} - y_i)], \tag{2.3.26}$$

where $y_{i+1}$ and $y_i$ are the (ordered) rapidities of the emissions connected by the propagator $t_i$. When doing dimensional regularisation, in $D = 4 - 2\varepsilon$ dimensions,

$$\hat{\alpha}(q_i) = -\frac{2g_s^2 C_A}{(4\pi)^{2+\varepsilon}} \frac{\Gamma(1-\varepsilon)}{\varepsilon}\left(\frac{q_{i\perp}^2}{\mu^2}\right)^\varepsilon \tag{2.3.27}$$

contains infrared divergences as $\varepsilon \to 0$. This prescription correctly describes the leading and even next-to-leading logarithmic terms, which has been verified up to N$^2$LO [112].

In eq. (2.2.4) we already saw the limit in which gluon $i$ became soft. Equivalent to eq. (2.2.5) we can explicitly integrate over the soft region $p_{i\perp} < \lambda$, with an IR cut-off $\lambda$,

$$\int_{y_{i-1}}^{y_{i+1}} \frac{\mathrm{d}y_i}{4\pi} \int_0^\lambda \frac{\mathrm{d}^{2+2\varepsilon} p_{i\perp}}{(2\pi)^{2+2\varepsilon}\,\mu^{2\varepsilon}} \frac{4g_s^2 C_A}{p_{i\perp}^2} = \frac{4g_s^2 C_A}{(4\pi)^{2+\varepsilon}} \Delta y_{i-1,i+1} \frac{1}{\varepsilon\Gamma(1+\varepsilon)}\left(\frac{\lambda^2}{\mu^2}\right)^\varepsilon. \tag{2.3.28}$$

Plugging in eq. (2.3.26) in eq. (2.3.5) gives a factor $\exp(2\Delta y\hat{\alpha})$. Expanding to first order in $\alpha_s$ and adding eq. (2.3.28) the $\varepsilon$-pole cancels

$$\frac{4g_s^2 C_A}{(4\pi)^{2+\varepsilon}}\Delta y_{i-1,i+1}\frac{1}{\varepsilon\mu^{2\varepsilon}}\left[\lambda^{2\varepsilon}\frac{1}{\Gamma(1+\varepsilon)} - q_{i\perp}^{2\varepsilon}\Gamma(1-\varepsilon)\right] \to -\Delta y_{i-1,i+1}C_A\frac{\alpha_s}{\pi}\ln\frac{q_{i\perp}^2}{\lambda^2}\,. \tag{2.3.29}$$

Equivalent to the general subtraction procedure in fixed order eq. (1.4.12), we can explicitly subtract the pole from the real

$$V_L^\mu V_{L\mu} \to V_L^\mu V_{L\mu} - \frac{4}{p_\perp^2}\theta\left(\lambda - p_\perp\right), \tag{2.3.30}$$

and add it to the virtual term. The remaining virtual correction is then similar to eq. (2.3.26)

$$\frac{1}{t_i} \to \frac{1}{t_i}\exp[\omega^0(q_{i\perp})(y_{i+1}-y_i)], \tag{2.3.31}$$

but now it only includes the finite remainder

$$\omega^0(q_{i\perp}) = -C_A\frac{\alpha_s}{\pi}\ln\frac{q_{i\perp}^2}{\lambda^2}\,. \tag{2.3.32}$$

For $p_\perp$ below 0.1 GeV eq. (2.3.30) can become numerically unstable. In practice integrating down to $p_\perp > \kappa = 0.2\,\text{GeV}$ is already enough to cancel the Lipatov vertex with its subtraction [98]. Furthermore if we choose to only subtract terms in the region of analytical integration, i.e. setting $\lambda = \kappa$, the matrix element is always positive.[4] Thus the integration does not need numerical cancellations, which speeds up convergence.

To conclude this section we collect all pieces and write down the full HEJ matrix element

$$\begin{aligned}\overline{\left|\mathcal{M}^{\text{HEJ}}_{f_a f_b\to f_1 g\cdots g f_n}\right|}^2 =& \frac{1}{4\,(N_C^2-1)}\,\left\|\mathcal{S}_{f_a f_b\to f_1 f_n}\right\|^2\left(g_s^2\;K_{f_1}\;\frac{1}{t_1}\right)\left(g_s^2\;K_{f_n}\;\frac{1}{t_{n-1}}\right)\\ &\cdot\prod_{i=1}^{n-2}\left[\frac{-g_s^2 C_A}{t_i t_{i+1}}\;\left(V_L^\mu(q_i,q_{i+1})V_{L\mu}(q_i,q_{i+1}) - \frac{4}{p_{i\perp}^2}\theta\left(\lambda-p_{i\perp}\right)\right)\right]\\ &\cdot\prod_{j=1}^{n-1}\exp\left[\omega^0(q_{j\perp})(y_{j+1}-y_j)\right].\end{aligned} \tag{2.3.33}$$

As promised we clearly see the different building pieces. All low multiplicity terms are hidden inside the currents $\mathcal{S}_{f_a f_b\to f_1 f_n}$, from which higher orders can be build up.

[4] In the following we will implicitly choose this convention without specifying the subtraction term eq. (2.3.30) at all.

## 2.4 Monte Carlo methods

Thus far we constructed matrix elements without worrying about the phase-space. In performing the phase-space integrals is rarely ever possible analytically, and instead one resorts to numerical methods. We will review the most common integration techniques for high energy physics in the following section. Typically the dimension of the phase space is high—for $n$ outgoing particles it has $D = 3n - 2$ dimensions[5]—while also including complex cuts from observables, IR regularisations, jet clustering, detector geometry, etc.

The easiest numerical integration is a more or less sophisticated slicing of the integration region into bins. The height of each bin is then set according to the function value (somewhere) in that bin. Thus the integral is approximated by a step function, with a trivial integral corresponding to total volume of all bins (Riemann sum). The shape of the bins can be further optimised to increase convergence, e.g. as a trapezoid or parabola (Simpson's rule). In all of these cases the total number of function calls is proportional to the total number of bins $N = M^D$, where $M$ is the number of bins in one dimension. Increasing the accuracy would require adding more bins. Hence the runtime of such algorithm scales exponentially with the number of dimensions, which becomes quickly unfeasible for higher dimensions. For example when using Simpson's rule the error scale like $M^{-4} = N^{-4/D}$.

In particle physics almost always random number based, so called *Monte Carlo*, integrations are used. Instead of slicing the phase space, $N$ random points $r_i$ are picked uniformly from anywhere in the integration region $\Omega$. The target function $f$ is then evaluated at these points. The mean value gives an estimate for the integral

$$E_N = \frac{1}{N}\sum_{i=0}^{N} f(r_i) \xrightarrow{N\to\infty} \langle f \rangle = \frac{1}{\mathrm{Vol}(\Omega)} \int_\Omega \mathrm{d}x f(x), \tag{2.4.1}$$

where $\mathrm{Vol}(\Omega)$ is the volume of $\Omega$. One can estimate the sampling uncertainty $s_N$ through the variance of the sampling

$$s_N^2 = \frac{V_N}{N-1} = \frac{1}{N(N-1)} \sum_{i=1}^{N} \left(f(r_i) - E_N\right)^2 \tag{2.4.2}$$

$$\xrightarrow{N\to\infty} \frac{\langle f^2 \rangle - \langle f \rangle^2}{N} = \frac{V(f)}{N}\,. \tag{2.4.3}$$

The uncertainty falls with every new point like $1/\sqrt{N}$, independent of the dimension-

[5]For each outgoing particle we integrate over their four-momenta, while keeping them on-shell. Additionally we have one overall energy-momentum conserving delta function, and for hadron colliders we also integrate over the energy fraction of the incoming partons $x_{a/b}$. Thus in total we integrate over $D = (4-1)n - 4 + 2 = 3n - 2$ dimensions.

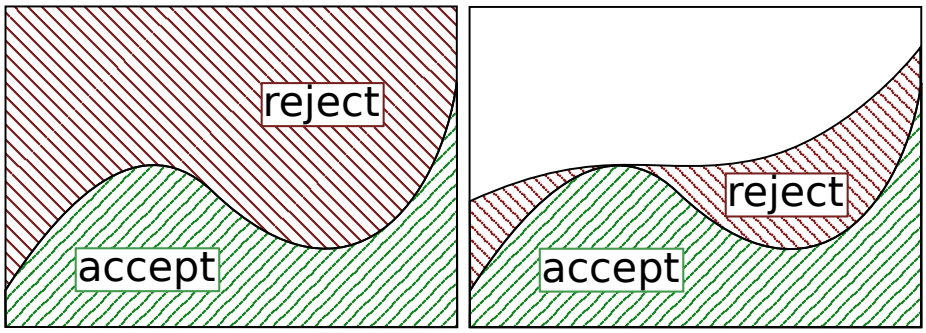



Figure 2.9: Illustration of hit-or-miss integration. Every point inside the upper, red region is rejected, and every point in the lower, green is accepted. For the naive sampling on the left random point are taken from the full integration region, while with a smarter over-estimator on the right no points are generated in the white region.

ality $D$. Hence when integrating more then 8 dimensions (or 3 final-states particles) Monte Carlo methods scale better with the number of function calls then Simpson's rule.

A concrete example of Monte Carlo implementation is the hit-or-miss technique. Both $x_i \in \Omega$ and $y_i$ are randomly selected, where $\min_x f(x) \leq y_i \leq \max_x f(x)$. If for a given $x_i$ the corresponding $y_i \leq f(x_i)$, we accept the point (hit), else we reject it (miss) (see left of fig. 2.9). The ratio of hits to misses is then the estimator for the integral

$$\mathrm{Vol}(\Omega)\left[\max_{x\in\Omega} f(x) - \min_{x\in\Omega} f(x)\right]\frac{N_{\mathrm{hit}}}{N_{\mathrm{total}}} \xrightarrow{N\to\infty} \int_\Omega \mathrm{d}x f(x). \tag{2.4.4}$$

Mathematically hit-or-miss is equivalent to the general method eq. (2.4.1) for the auxiliary function

$$\tilde{f}(x,y) = \begin{cases} 1 & f(x) \leq y \\ 0 & \text{otherwise} \end{cases}. \tag{2.4.5}$$

To further improve the sampling one has to reduce the variance of $f$ in eq. (2.4.3). In the hit-or-miss scenario one tries to avoid misses, by using a suitable over-estimator $g(x)$ and only selecting $y_i$ in the range $[\min_x f(x), g(x_i)]$ (see right of fig. 2.9). The total integration volume changes from $\mathrm{Vol}(\Omega)$ to $\mathrm{Vol}(\Omega_g) = \int \mathrm{d}x g(x)$, which has to be known analytically. The improvement is therefore limited by the amount of a priori information of $f$ put into $g$.

Instead of only limiting the generation of $y$, *importance sampling* more generally biases the $x$-distribution directly, by selecting $x$ according to a distribution $G = \int \mathrm{d}x g$.

Formally this is a variable transformation

$$\int \mathrm{d}x f(x) = \int \mathrm{d}x \frac{\mathrm{d}G(x)}{\mathrm{d}x} \frac{f(x)}{\frac{\mathrm{d}G(x)}{\mathrm{d}x}} = \int \mathrm{d}G(x) \frac{f(x)}{g(x)} \,. \tag{2.4.6}$$

This introduces a weight $1/g$ to undo the sampling bias.[6] The resulting variance is then reduced to the variance of $f/g$. Again the closer $g$ mimics $f$ the quicker the convergence, with the obvious perfect case of zero variance at $g = f$.

We need to generate $x$ according to $G$. Hence, $g$ has to be a probability density, i.e. any function with $g(x) \geq 0, \forall x \in \Omega$ and $\int_\Omega \mathrm{d}x g(x) = 1$. It is no longer required that $g(x) \geq f(x)$. But if $g \to 0$ for $f \neq 0$ this region will be under-sampled. With finite statistics under-sampling can mean that some strongly suppressed values are missed, even if $g(x) > 0$. This can create artificial holes in the phase space, resulting in a wrong value without an increased variance as indication. In any case under-sampling, as in general any "bad" sampling, will increase the variance. Making incorrect assumptions about $f$ can be worse then making no assumptions at all.

Additionally we have to be able to generate $x$ according to $G$, either directly through a (black box) pseudo-random number generator, or by converting a flat distribution of $r$ to a $g$-distribution of $x$. For the latter we use the ansatz

$$\int_{x_\mathrm{min}}^{x} \mathrm{d}x' g(x') = r \int_{x_\mathrm{min}}^{x_\mathrm{max}} \mathrm{d}x' g(x'), \tag{2.4.7}$$

which we can solve for $x$

$$x = G^{-1} \left( r \left[ G(x_\mathrm{max}) - G(x_\mathrm{min}) \right] + G(x_\mathrm{min}) \right). \tag{2.4.8}$$

Thus we have to know the inverse $G^{-1}$, which limits the choice of $g$ even further.

One trick to find a suitable $g$ is to build it up from multiple channels

$$g(x) = \sum_i n_i g_i(x), \tag{2.4.9}$$

where each $g_i$ catches some characteristic of $f$. Each of these functions is then weighted with some relative weight $n_i > 0$, such that $\sum n_i = 1$. For example in QCD each $g_i$ could account for one IR pole. One can also use a step function for $g_i$, similar to the slicing in a Riemann sum. In particular, it is possible to adjust the steps automatically to reduce the variance, without prior knowledge of $f$, as done in the VEGAS algorithm [113].

[6] Note that for multi-dimensional integrals $1/g$ is the Jacobian $\det(\mathrm{d}x/\mathrm{d}G)$.

# Chapter 3

# Matching HEJ with fixed order

The standard procedure to suppress Higgs production through gluon fusion compared to that from weak-boson fusion is to cut away the low energetic gluons. However the impact of the radiative corrections to each process is rather different; in particular, the $t$-channel colour octet exchange in the gluon-fusion process leads to increased jet-activity [114], which allows for a distinction of the production mechanism within the phase space populated by weak boson fusion. The two jets in weak boson fusion are often separated by a large invariant mass and rapidity span. This is the phase-space region where the perturbative corrections for the QCD processes contain logarithms of $\log|\hat{s}/\hat{t}|$ from BFKL [22–25]. As we saw before these logarithms are contained within the formalism of High Energy Jets, where a systematic treatment is obtained by a power-expansion of the scattering matrix element in $\hat{s}/\hat{t}$. The first sub-leading corrections were presented for the production of a Higgs boson with dijets in ref. 98 by calculating the leading behaviour of certain sub-leading processes. This constitutes control of a well-defined set of NLL BFKL logarithms within HEJ. These logarithms drive the pattern of further emissions from the QCD process [114], which allows for a better discrimination between the GF and VBF processes than what could be performed by investigating the dynamics of just two jets in the event.

As discussed in section 2.3, the formalism of HEJ captures leading logarithmic terms to processes with at least two jets at large partonic centre-of-mass energy of the form $\alpha_s^k \log|\hat{s}/\hat{t}|^k \simeq \alpha_s^k \Delta y_{j_f j_b}^k$, where $\Delta y_{j_f j_b}$ is the rapidity-difference between the jets forward and backward in rapidity. The systematic treatment of these terms is based on a logarithmic all-order expansion point-by-point in phase space of the leading virtual corrections to all orders, combined with a power-expansion in $\hat{s}/\hat{t}$ of the square of the tree-level amplitudes, again point-by-point in the $n$-particle phase space. Upon integration, the leading power-expansion of the square of the amplitudes ensure the appropriate logarithmic accuracy of cross sections. Through

the numerical integration HEJ generates fully differential events, allowing for detailed jet clustering and event analyses.

Within the formalism of HEJ, the $m$-jet rates entering each prediction are matched to LO accuracy point-by-point in phase space by the following procedure for mapping the $n$-parton resummation phase space point into a $m$-parton tree-level phase space point, described in more detail in ref. 34 and section 3.1:

1. cluster the $n$-parton phase space point into jets with a chosen jet algorithm and jet $p_\perp$-threshold (e.g. anti-$k_t$ clustering, with a threshold of 30 GeV),

2. remove the partons not clustered into the jets from the event, and distribute the sum of their transverse momenta onto the hard jets,

3. adjust the energy and longitudinal, momentum of each jet such that it is on-shell, while keeping their rapidities fixed

4. adjust the momenta of the incoming partons such that energy and momentum conservation is restored.

The result of this procedure is a set of momenta for which the on-shell $m$-jet tree-level matrix element can be evaluated. This in turn allows for the weight of the generated event to be reweighted to full tree-level $m$-jet accuracy, thus obtaining full tree-level accuracy up to the multiplicity for which the tree-level matrix elements can be evaluated in reasonable time. This method for matching the all-order results to fixed-order high-multiplicity matrix elements has been used for matching all results obtained with HEJ: jets, and $\gamma/Z$, $W$ plus at least two jets with matching up to 4 jets, and $H$ with at least two jets, with matching up to 3 jets.

The matching procedure described here can thus be viewed as *merging* the results of fixed-order calculations by use of the power-expanded matrix elements of HEJ coupled with the logarithmic virtual corrections, similar to the CKKW-L–method [115, 116] of using the logarithmic accuracy of a shower-algorithm to merge fixed-order cross sections of varying multiplicity. This chapter will present a complete reformulation of the procedure for merging and all-order summation. With the same input (as in use of the same matrix elements to the same order), the results are unchanged, but the new procedure for obtaining the all-order results and the merging will allow for merging results beyond tree-level, whilst being computationally much more efficient. Additionally, in the future the algorithm presented here can potentially be generalised to NLO to give LL + NLO correct results, leading to reduced scale variations.

In section 3.1 we describe the original mechanism for matching leading-order samples within HEJ before a detailed discussion of the new formulation. This includes both analytical aspects and practical aspects of implementation. In section 3.2 we study the results obtained in the new formalism in the context of Higgs boson plus dijets in three studies. Firstly we confirm that matching to fixed-order samples is limited to a maximum of three jets, we find consistent results with the previous formalism. Secondly, we study the impact of increasing the multiplicity in the fixed-order samples, now possible for the first time. Thirdly, we compare the matched all-order results of HEJ with those obtained at next-to-leading order accuracy. We conclude in section section 3.3 with a final discussion. The matching presented here is publicly available in HEJ version 2.

## 3.1 Matching

In the original formulation, the cross sections within HEJ are calculated by explicitly constructing the all-order result. We therefore first generated a $2 \to n + l$ kinematic point for each number of partons $n = 2, \ldots, N$, where $N$ is chosen sufficiently large (in practice around 22), and $l$ describes the non-partonic particles produced, e.g. $Z$, $W$, $H$ or their decay products. In order to simplify the notation we will only discuss the purely partonic case $l = 0$. Likewise, we will restrict our discussion to the leading-logarithmic contribution. Note that all our arguments apply equally to the more general scenario. We demonstrate this by showing results for the production of a Higgs boson in association with at least two jets, including computed sub-leading corrections [98].

The high-energy limit is dominated by FKL configurations, where two partons scatter in such a way that there is no radiation outside the rapidity range spanned by the scattering partons and only gluons are emitted inside this range. To ensure the high-energy limit applied is valid, it is required that the extremal (in rapidity) partons are perturbative (hard in terms of transverse momentum), and are members of the extremal jets. The transverse momenta of the remaining partons are all generated down to effectively 0 GeV. The matching to LO accuracy for all $m$-jet rates, $m \leq n$, is then obtained by first projecting the kinematics of the generated all-order events into Born kinematics according to the number of hard jets as described in the previous section. The event weight is multiplied with a ratio of the square of the full Born-level amplitude to the HEJ approximation of the same. The cross section

(and kinematic distributions) is then obtained through the formula[1]

$$\begin{aligned}\sigma_{2j}^{\text{resum,match}} = \sum_{f_1,f_2} \sum_{n=2}^{\infty} \int_{p_{1\perp}=p_{\min\perp}}^{p_{1\perp}=\infty} \frac{\mathrm{d}^2\mathbf{p}_{1\perp}}{(2\pi)^3} \int_{p_{n\perp}=p_{\min\perp}}^{p_{n\perp}=\infty} \frac{\mathrm{d}^2\mathbf{p}_{n\perp}}{(2\pi)^3} \prod_{i=2}^{n-1} \int_{p_{i\perp}=\lambda}^{p_{i\perp}=\infty} \frac{\mathrm{d}^2\mathbf{p}_{i\perp}}{(2\pi)^3} \\ \times \ \mathbf{T}_y \prod_{i=1}^{n} \left( \int \frac{\mathrm{d}y_i}{2} \right) \ \times \frac{\overline{\left|\mathcal{M}_{\text{HEJ}}^{\text{reg}}(\{p_i\})\right|^2}}{\hat{s}^2} \times \sum_m \mathcal{O}_{mj}^e(\{p_i\}) \ w_{m-\text{jet}}^{\text{LO}} \\ \times \ \ x_a f_{a,f_1}(x_a,Q_a) \ x_b f_{b,f_2}(x_b,Q_b) \ (2\pi)^4 \ \delta^2\left(\sum_{i=1}^{n} \mathbf{p}_{i\perp}\right),\end{aligned} \tag{3.1.1}$$

where $|\mathcal{M}_{\text{HEJ}}^{\text{reg}}(\{p_i\})|^2$ is the square of the regularised all-order matrix element within HEJ for the $2 \to n$ phase space point, and

$$w_{m-\text{jet}}^{\text{LO}} \equiv \frac{\overline{\left|\mathcal{M}_{\text{LO}}^{f_1 f_2 \to f_1 g \ldots g f_2}\left(\left\{p_{\mathcal{J}_l}^B(\{p_i\})\right\}\right)\right|^2}}{\left|\mathcal{M}_{\text{LO,HEJ}}^{f_1 f_2 \to f_1 g \ldots g f_2}\left(\left\{p_{\mathcal{J}_l}^B(\{p_i\})\right\}\right)\right|^2} \tag{3.1.2}$$

is the ratio between the matrix element evaluated at full tree-level accuracy and within HEJ for the state projected to tree-level $2 \to m$ kinematics described by the jet momenta $\left\{p_{\mathcal{J}_l}^B(\{p_i\})\right\}$. $\mathbf{T}_y$ indicates rapidity ordering, while $\mathcal{O}_{mj}^e(\{p_i\})$ is the exclusive $m$-jet measure applied to the generated event kinematics, i.e.

$$\mathcal{O}_{mj}^e(\{p_i\}) = \begin{cases} 1 & \text{all particles } \{p_i\} \text{ are clustered into } \textit{exactly } m \text{ jets} \\ 0 & \text{otherwise} \end{cases} \tag{3.1.3}$$

selects the number of jets $m$ for each momenta configuration $\{p_i\}$.

The limits of the integral over the transverse momentum of the extremal partons combined with the two-jet measure is set to guarantee that the extremal partons carry the dominant momentum of the extremal jets. We choose a cut-off $p_{\min\perp}$ corresponding to 90% of the transverse momentum of the respective extremal jet.[2] We use here the phrase 'kinematics of the generated all-order event' to mean the $n$-parton kinematic point of the resummation event sampled in eq. (3.1.1). In order to match each $m$-jet rate to tree-level accuracy, each generated event in the all-order phase-space is mapped to a $m$-jet tree-level kinematic point, where each jet contains exactly one parton, and requires an evaluation of the full $m$-parton matrix element.

The scale-variation of the normalisation of the cross sections is determined by the LO matrix elements, and mostly unchanged by the leading logarithmic high-energy

[1] To not clutter the notation, we omit the integrals over the incoming kinematics $x_{a/b}$ and $Q_{a/b}$.

[2] This is a slightly more sophisticated cut than that investigated in ref. 34, 98, 107, 108, and ensures that the soft divergence which would be regulated at next-to-leading logarithmic accuracy of the extremal currents does not impact the result obtained with the leading-logarithmic currents even for jets at large transverse momentum.

resummation implemented in HEJ. This could be reduced by extending the reweighting factor $w_{m-\text{jet}}$ to next-to-leading order accuracy. However, in order to do this, one would have to integrate over all $m+1$ parton real emission phase space resulting in a specific $m$-jet Born-level kinematics. This would be prohibitively time-consuming. An alternative approach is to begin with fixed-order samples of exclusive jet rates and then merge these using HEJ to generate all-order results. We demonstrate how to do this in the next subsection and find significant benefits even at tree level. In particular, each phase space point used for the tree-level matrix element maps into all the relevant resummation phase space points which leads to fewer evaluations of the tree-level matrix elements. This in turn allows for matching to higher multiplicity with a given CPU envelope.

### 3.1.1 Supplementing fixed order samples with HEJ resummation

The reformulation of the resummation and matching should reproduce the results of eq. (3.1.1). Starting from this equation, we introduce a $\delta$-functional and an integration over the Born-level kinematics of the on-shell, reshuffled jets $\{j_B^i\}$ reconstructed from the resummed kinematics. Equation (3.1.1) is rewritten to

$$
\begin{aligned}
\sigma_{2j}^{\text{resum,match}} = \sum_{f_1,f_2} \sum_m \prod_{j=1}^{m} \left( \int_{p_{j\perp}^B=0}^{p_{j\perp}^B=\infty} \frac{\mathrm{d}^2\mathbf{p}_{j\perp}^B}{(2\pi)^3} \int \frac{\mathrm{d}y_j^B}{2} \right) (2\pi)^4 \, \delta^{(2)}\left( \sum_{k=1}^{m} \mathbf{p}_{k\perp}^B \right) \\
\times \; x_a^B f_{a,f_1}(x_a^B, Q_a^B) \; x_b^B f_{b,f_2}(x_b^B, Q_b^B) \; \frac{\overline{\left|\mathcal{M}_{\text{LO}}^{f_1 f_2 \to f_1 g \ldots g f_2}\left(\left\{p_j^B\right\}\right)\right|^2}}{\left(\hat{s}^B\right)^2} \\
\times \frac{w_{m-\text{jet}}}{\overline{\left|\mathcal{M}_{\text{LO}}^{f_1 f_2 \to f_1 g \ldots g f_2}\left(\left\{p_j^B\right\}\right)\right|^2}} \; (2\pi)^{-4+3m} \; 2^m \\
\times \sum_{n=2}^{\infty} \int_{p_{1\perp}=p_{\min\perp}}^{p_{1\perp}=\infty} \frac{\mathrm{d}^2\mathbf{p}_{1\perp}}{(2\pi)^3} \int_{p_{n\perp}=p_{\min\perp}}^{p_{n\perp}=\infty} \frac{\mathrm{d}^2\mathbf{p}_{n\perp}}{(2\pi)^3} \\
\times \prod_{i=2}^{n-1} \left( \int_{p_{i\perp}=\lambda}^{p_{i\perp}=\infty} \frac{\mathrm{d}^2\mathbf{p}_{i\perp}}{(2\pi)^3} \right) \mathbf{T}_y \prod_{i=1}^{n} \left( \int \frac{\mathrm{d}y_i}{2} \right) (2\pi)^4 \, \delta^{(2)}\left( \sum_{k=1}^{n} \mathbf{p}_{k\perp} \right) \\
\times x_a f_{a,f_1}(x_a, Q_a) \; x_b f_{b,f_2}(x_b, Q_b) \; \frac{\overline{\left|\mathcal{M}_{\text{HEJ}}^{f_1 f_2 \to f_1 g \ldots g f_2}(\{p_i\})\right|^2}}{\hat{s}^2} \\
\times \mathcal{O}_{mj}^e \left( \prod_{l=1}^{m-1} \delta^{(2)}(\mathbf{p}_{\mathcal{J}_l\perp}^B - \mathbf{j}_{l\perp}) \right) \left( \prod_{l=1}^{m} \delta(y_{\mathcal{J}_l}^B - y_{\mathcal{J}_l}) \right) \\
\times \frac{\left(\hat{s}^B\right)^2}{x_a^B f_{a,f_1}(x_a^B, Q_a^B) \; x_b^B f_{b,f_2}(x_b^B, Q_b^B)} \, .
\end{aligned}
\tag{3.1.4}
$$

The first two lines are now the phase space integration over the LO matrix element, which can be represented in terms of (potentially weighted) tree-level events. Obviously, the Born-level partonic momenta are identical to the Born-level jet momenta, i.e. $p_i^B \equiv p_{\mathcal{J}_i}^B$, so that

$$\frac{w_{m-\text{jet}}}{\overline{\left|\mathcal{M}_{\text{LO}}^{f_1 f_2 \to f_1 g \ldots g f_2}\left(\left\{p_j^B\right\}\right)\right|^2}} = \overline{\left|\mathcal{M}_{\text{LO,HEJ}}^{f_1 f_2 \to f_1 g \ldots g f_2}\left(\left\{p_j^B\right\}\right)\right|}^{-2} \tag{3.1.5}$$

only depends on the Born-level HEJ approximation to the matrix element. Lines 4 to 6 preform the integration of the HEJ matrix elements over all of the resummation phase space. The last line removes the factors introduced in the first line of eq. (3.1.4) compared to eq. (3.1.1) in order to write the matching in terms of a standard phase space integration over fixed-order PDFs and matrix elements.

The $\delta$-functionals of the second of last line in eq. (3.1.4) connect the reconstructed Born-level kinematics with the kinematics of the jets arising from the resummation. The algorithm devised for projecting the jet momenta of the resummation onto Born-level kinematics gives [34]

$$\mathbf{p}_{\mathcal{J}_{l\perp}}^B = \mathbf{j}_{l\perp} \equiv \mathbf{p}_{\mathcal{J}_l\perp} + \mathbf{q}_\perp \frac{|\mathbf{p}_{\mathcal{J}_l\perp}|}{P_\perp}\,, \tag{3.1.6}$$

plus the constraint that the rapidities of the jets are kept fixed. Here $\mathbf{p}_{\mathcal{J}_l}^B$ is the momentum of the fixed-order, matching level jet, and $P_\perp$ is the scalar sum of the jet transverse momenta after resummation $P_\perp = \sum_{j=1}^m |\mathbf{p}_{\mathcal{J}_j\perp}|$. The sum of the transverse momenta of partons outside jets $\mathbf{q}_\perp$ equals minus the transverse momentum of the jets after resummation $\mathbf{q}_\perp = -\sum_{j=1}^m \mathbf{p}_{\mathcal{J}_j\perp}$.

This algorithm can be straightforwardly applied when the resummation event has been constructed, and had a jet-clustering applied. If, however, we want to start from fixed-order generated events, the algorithm needs to be inverted, such that all resummation-momenta on the right-hand side of eq. (3.1.6) are explored for a given Born-level kinematic point. This inversion has to be done numerically, which makes it computationally expensive. To circumvent this one can construct an "alternative reshuffling"

$$\mathbf{p}_{\mathcal{J}_{l\perp}} = \mathbf{j}_{l\perp}^B \equiv \mathbf{p}_{\mathcal{J}_l\perp}^B - \mathbf{q}_\perp \frac{|\mathbf{p}_{\mathcal{J}_l\perp}^B|}{P_\perp^B}\,, \tag{3.1.7}$$

which is linear in the resummation momenta.[3] The difference between both mappings is subleading, and in practice negligibly small. To match the previous implementation

[3] This corresponds to a variable transformation $\mathbf{p}_{\mathcal{J}_{l\perp}}^B \to \mathbf{p}_{\mathcal{J}_{l\perp}}$ in eq. (3.1.4). Thus we get an extra Jacobian $\left|\partial \mathbf{j}_{l'\perp}^B / \partial \mathbf{p}_{\mathcal{J}_l\perp}^B\right|$.

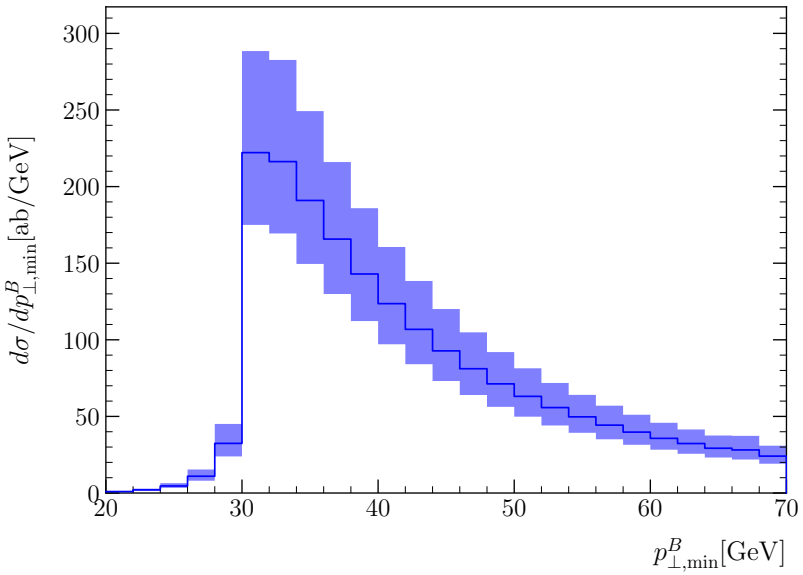



Figure 3.1: Distribution of the minimal transverse energy of the jets *before* resummation, $p^B_{\mathrm{min}\perp}$, contributing for generating jets *after* resummation with HEJ for Higgs boson plus dijet production. After resummation we require a jet transverse momentum of at least 30 GeV, below this scale the distribution falls off quickly. Hence the resummation phase space with a given minimum jet transverse momentum is covered by a fixed-order generation with a slightly smaller requirement on the jet transverse momentum. For example, a generation of fixed-order events with a minimum jet transverse momentum of 20 GeV is sufficient for an analysis requiring a transverse momentum of at least 30 GeV.

of HEJ from ref. 34, we use eq. (3.1.6) for this chapter. The numerical inversion is done by GLS routines [117]. Contrary, in chapter 4 we choose eq. (3.1.7), which also became the new default in HEJ 2.

While eq. (3.1.4) is mathematically equivalent to eq. (3.1.1), it does not prove that the approach is viable. The first challenge is to ensure that in fact, the integration over the matching, or fixed-order phase space, in the first line of eq. (3.1.4) does not actually extend to zero transverse momentum of the matching jets. This would lead to a divergence in the fixed-order cross section and invalidate the whole procedure. In fig. 3.1 we investigate the minimum transverse momentum of jets used in the matching for the evaluation of fixed-order matrix elements. The plot shows $\mathrm{d}\sigma/\mathrm{d}p^B_{\mathrm{min}\perp}$, where $p^B_{\mathrm{min}\perp}$ is the minimum jet transverse momentum used in the merging with matrix elements—i.e. the minimum transverse momentum in the resulting on-shell Born-level kinematics after reshuffling—for Higgs-boson production in association with at least two jets with transverse momentum of at least 30 GeV.

One sees that the matrix element sample needs to include events with a minimum jet transverse momentum *below* the final analysis scale, but not "too far" below. It is observed that this distribution gets broader, and that the weight for small $p^B_{\text{min}\perp}$ is relatively important, both for larger rapidity spans, and if more hard jets are required (obviously these two requirements are linked).

The next challenge is to generate all resummation kinematics corresponding to a specific fixed-order or matching kinematics. This is not an obvious switch to make: substituting a requirement on LO kinematics to result in a given Born-level jet configuration with that of the full resummation event resulting in a given Born-level jet kinematics. However, the formalism will offer a number of benefits. Statistical convergence can be controlled at a more fine-grained level. Stability can be ensured first at the fixed-order stage before attempting resummation, and each jet multiplicity can be considered separately. We are free to choose whichever generators we find most suitable for producing fixed-order events. A further improvement is that the fixed-order matrix element is evaluated only once for each fixed-order kinematic point, generating a significant enhancement to the computational efficiency, especially for high jet multiplicities.

### 3.1.2 Phase space generation

In order to perform the resummation, we are tasked with the numerical evaluation of the last five lines of eq. (3.1.4). In principle, we have to integrate over the phase space of arbitrarily many further real emissions. This is made feasible by the fact that for a given fixed-order configuration with finite rapidity span, only a limited number of additional gluons actually lead to a non-negligible contribution in the resummation. Still, the typical multiplicities in the interesting region of large rapidity separations will be quite high and we are required to inspect the corresponding high-dimensional phase space carefully for an efficient integration. In the following section, we discuss how to construct an efficient importance sampling.

#### Gluon multiplicity

The typical number of extra emissions depends strongly on the rapidity span of the underlying fixed-order event. Let us, for example, consider a fixed-order FKL-type multi-jet configuration with rapidities $y_{j_f}$, $y_{j_b}$ of the most forward and backward jets, respectively. By construction in the matching algorithm of ref. 34, the jet multiplicity and the rapidity of each jet are conserved when adding resummation. This implies that additional hard radiation is restricted to rapidities $y$ within a

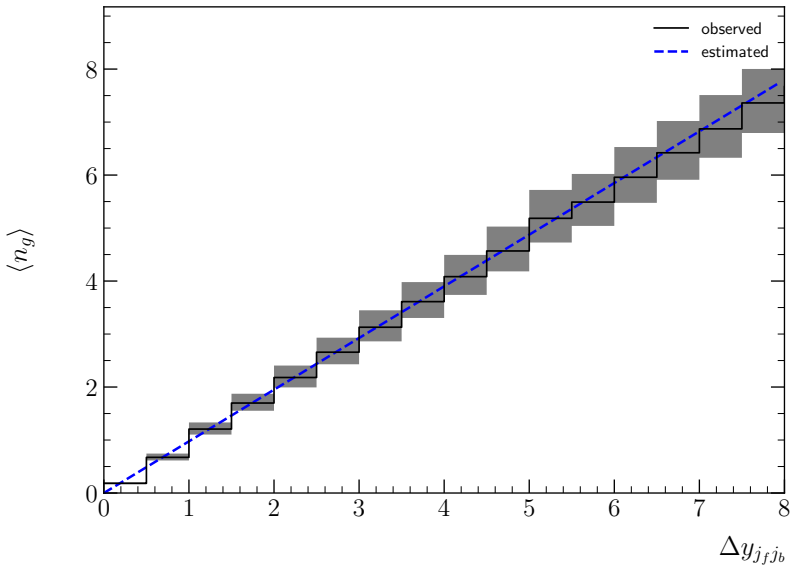



Figure 3.2: Average number of additional gluon emissions $n_g$ as a function of the rapidity span between the extremal jets. The black histogram shows the predicted numbers in the production of a Higgs boson in association with at least two jets. The blue, dashed line is the fitted function used for the phase space generation.

region $y_{j_b} \lesssim y \lesssim y_{j_f}$. Within HEJ, we require the most forward and most backward emissions to be hard in order to avoid divergences [34], so this constraint in fact applies to *all* additional radiation.

To simplify the remaining discussion, let us remove the FKL rapidity ordering

$$\mathbf{T}_y \prod_{i=1}^{n} \int \frac{\mathrm{d}y_i}{2} = \frac{1}{n!} \prod_{i=1}^{n} \int \frac{\mathrm{d}y_i}{2}\,, \tag{3.1.8}$$

where all rapidity integrals now cover a region which is approximately bounded by $y_{j_b}$ and $y_{j_f}$. Each of the $m$ jets has to contain at least one parton; selecting random emissions we can rewrite the phase space integrals as

$$\frac{1}{n!} \prod_{i=1}^{n} \int [\mathrm{d}p_i] = \left( \prod_{i=1}^{m} \int [\mathrm{d}p_i]\ \mathcal{J}_i(p_i) \right) \frac{1}{n_g!} \prod_{i=m+1}^{m+n_g} \int [\mathrm{d}p_i] \tag{3.1.9}$$

with jet selection functions

$$\mathcal{J}_i(p) = \begin{cases} 1 & p \text{ clustered into jet } i \\ 0 & \text{otherwise} \end{cases} \tag{3.1.10}$$

and the number of extra gluon emission $n_g \equiv n - m$. Here and in the following we use the short-hand notation $[\mathrm{d}p_i]$ to denote the phase-space measure for parton $i$. As

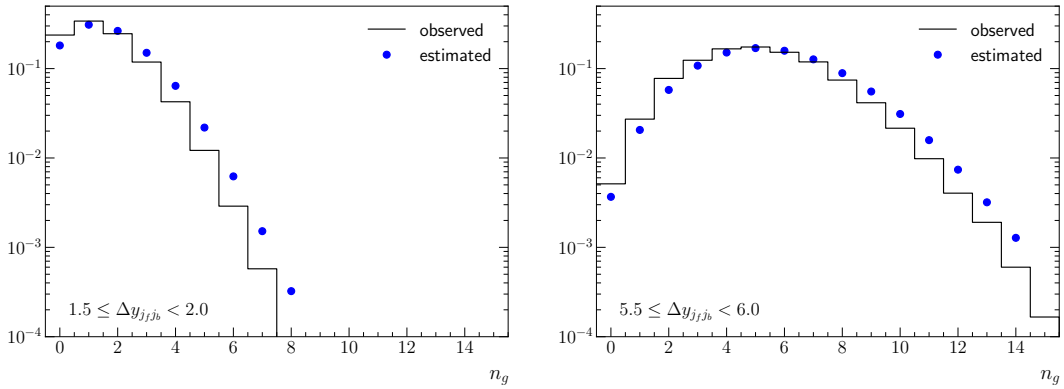



Figure 3.3: Number of additional gluon emissions $n_g$ for two different rapidity spans between the extremal jets for $\Delta y_{j_f j_b} = 1.75$ (left) and $\Delta y_{j_f j_b} = 5.75$ (right). The estimates in blue are based on Poisson distributions with mean values taken from the fit function in fig. 3.2, and in black are the observed values after integration.

is evident from eq. (3.1.9), adding an extra emission $n_g + 1$ introduces a suppression factor $\frac{1}{n_g+1}$. However, the additional phase space integral also results in an enhancement proportional to $\Delta y_{j_f j_b} = y_{j_f} - y_{j_b}$. This is a result of the rapidity-independence of the MRK limit of the integrand, consisting of the matrix elements divided by the flux factor. Indeed, we observe that the typical number of gluon emissions is to a good approximation proportional to the rapidity separation and the phase space integral is dominated by events with $n_g \approx \Delta y_{j_f j_b}$ (see fig. 3.2).

For the actual phase space sampling, we assume a Poisson distribution and extract the mean number of gluon emissions in different rapidity bins and fit the results to a linear function in $\Delta y_{j_f j_b}$, finding a coefficient of 0.975 for the inclusive production of a Higgs boson with two jets. In figs. 3.2 and 3.3 we compare the fit with the actual outcome.

### Number of gluons inside jets

For each of the $n_g$ gluon emissions we can split the phase-space integral into a (disconnected) region inside the jets and a remainder

$$\int [\mathrm{d}p_i] = \int [\mathrm{d}p_i]\, \theta\left(\sum_{j=1}^{m} \mathcal{J}_j(p_i)\right) + \int [\mathrm{d}p_i] \left[1 - \theta\left(\sum_{j=1}^{m} \mathcal{J}_j(p_i)\right)\right]. \tag{3.1.11}$$

We choose an importance sampling which is flat in the plane spanned by the azimuthal angle $\phi$ and the rapidity $y$. This is observed in BFKL and valid in the MRK limit. Furthermore, we assume (non-overlapping) anti-$k_t$ jets, which cover an area of $\pi R^2$ [70].

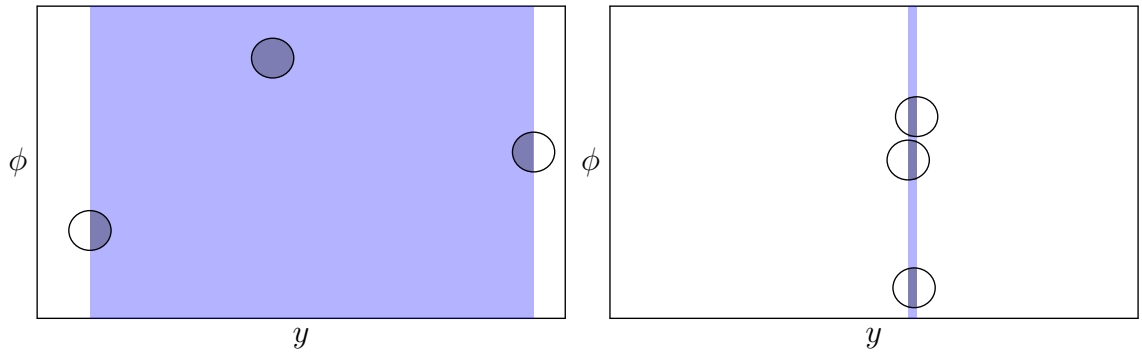



Figure 3.4: Estimated phase space areas for the emission of extra gluons for sample three-jet configurations. The left panel shows the case of a large rapidity separation. On the right we illustrate the estimate for a very small rapidity span.

In principle, the total accessible area in the $y$-$\phi$ plane is given by $2\pi\Delta y_{fb}$, where $\Delta y_{fb} \geq \Delta y_{j_f j_b}$ is the a priori unknown rapidity separation between the most forward and backward partons. In most cases the extremal jets consist of single partons, so that $\Delta y_{fb} = \Delta y_{j_f j_b}$. For the less common case of two partons forming a jet we observe a maximum distance of $R$ between the constituents and the jet centre. In rare cases jets have more than two constituents. Empirically, they are always within a distance of $\frac{5}{3}R$ to the centre of the jet [118], so $\Delta y_{fb} \leq \Delta y_{j_f j_b} + \frac{10}{3}R$. In practice, the extremal partons are required to carry a large fraction of the jet transverse momentum (cf. section 3.1) and will therefore be much closer to the jet axis.

In summary, for sufficiently large rapidity separations we can use the approximation $\Delta y_{fb} \approx \Delta y_{j_f j_b}$. If there is no overlap between jets, the probability $p_{\mathcal{J},>}$ for an extra gluon to end up inside a jet is then given by (cf. fig. 3.4)

$$p_{\mathcal{J},>} = \frac{(m-1)R^2}{2\Delta y_{j_f j_b}}\,. \tag{3.1.12}$$

For a very small rapidity separation, eq. (3.1.12) obviously overestimates the true probability. The maximum phase space covered by jets in the limit of a vanishing rapidity distance between all partons is $m \cdot 2R\Delta y_{fb}$. We therefore estimate the probability for a parton to end up inside a jet as

$$p_{\mathcal{J}} = \min\left(\frac{(m-1)R^2}{2\Delta y_{j_f j_b}}, \frac{mR}{\pi}\right). \tag{3.1.13}$$

In fig. 3.5 we compare this estimate with the actually observed fraction of additional emissions into jets. We observe good agreement over the entire rapidity range and for different jet multiplicities.

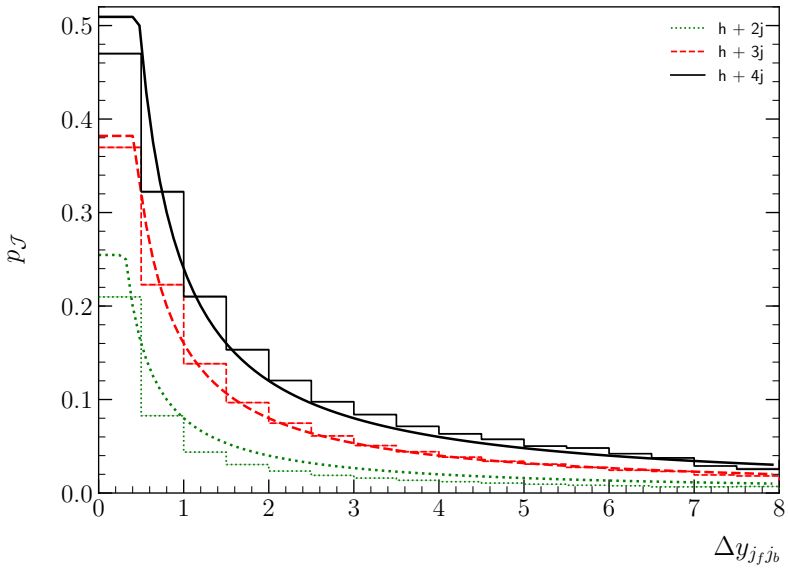



Figure 3.5: Estimated probability (lines) for an extra emission to end up inside a jet compared to the fraction observed (histograms) in the exclusive production of a Higgs boson with two (green, dotted), three (red, dashed), and four (black, solid) jets.

### Gluons outside jets and observed jet momenta

Using our estimate for the probability of a gluon to be a jet constituent, we arrive at a number $n_{g,\mathcal{J}}$ of gluons inside jets. Before integrating over their remaining phase space, we first have to determine the momenta $p_{\mathcal{J}_i}$ of the observed (resummation) jets from eq. (3.1.6). To this end, we have to determine the total transverse momentum $\mathbf{q}_\perp$ of the gluons *outside* jets. After generating soft transverse momenta for these $n_g - n_{g,\mathcal{J}}$ gluons, we invert the non-linear system eq. (3.1.6), or, in newer implementations, the linear eq. (3.1.7). Note that we have to postpone the rapidity integration, since at this point the rapidity span in the phase space integral is not yet known. The most forward and backward partons have to be part of the extremal jets. Therefore, their momenta will only be determined in the next step.

### Gluons inside jets

Recall that after the first step in the phase space parametrisation, eq. (3.1.9), each jet has exactly one constituent. We now assign each of the $n_{g,\mathcal{J}}$ gluons to a random jet. For jets with a single constituent, the parton momentum is fixed completely by the constraints in eq. (3.1.4). In the case of two constituents, we observe that the partons are always inside the jet cone with radius $R$ and often very close to the jet centre. This allows efficient integration by choosing a distance to the jet centre and

an azimuthal angle with respect to the jet axis for one of the partons as integration variables, which determines all momentum components of both constituents.

As is evident from fig. 3.5, jets with three or more constituents are rare and an efficient phase-space sampling is less important. For such jets, we exploit the observation that partons with a distance larger than $R_{\max} = \frac{5}{3}R$ [118] to the jet centre are never clustered into that jet. Assuming $N$ constituents, we choose distances, angles, and transverse momenta for $N-1$ of them and determine the momentum of the last constituent from the requirement that the constituent momenta have to add up to the jet momentum. Since this last momentum may lie outside the jet cone, it is mandatory to check explicitly whether all candidates are actually clustered into the considered jet. This is to ensure the correct coverage of phase space.

After constructing the resummation jets, we are now in the position to evaluate the rapidity integrals for the partons outside the jets. Finally, we use `FastJet` [69] to recluster all emitted partons into jets again to check whether the reshuffling conditions imposed by eq. (3.1.4) are fulfilled. We also ensure that all partons are assigned as intended, i.e. the $n_{g,\mathcal{J}}$ designated jet constituents are indeed part of their respective jet and all remaining partons end up outside jets.

### 3.1.3 Technical aspects

We have now outlined the practical steps necessary to numerically solve eq. (3.1.4). The complete algorithm is implemented in version 2 of the C++ program High Energy Jets, which was first publicly released in ref. 3 and can be downloaded from ref. 4, 5. Beside the complete rewrite of the integration, HEJ 2 also implemented the finite quark mass effects for Higgs-boson production, which we will study in chapter 4. The remaining processes, $Z/\gamma$ or $W$ boson with jets, from the old version and expanding them to include all NLL configurations [110] will be available with the next release.

One design goal for HEJ was to make it independent of any specific Monte Carlo generator. To achieve this, we use the *Les Houches event file* (LHEF) standard [119, 120] to read input events. Though LHEF is not ideal[4] it is the most widely used format in heigh energy physics and is supported by most generators. Additionally it is similarly used by Pythia as an input, which ensures at least some compatibility. To lift some of the restrictions from the XML based LHEF standard, with the next release HEJ will support the format proposed in ref. 121 based on HDF5 files [122], by using the header-only library HighFive [123]. HDF5 files are binary data files

[4] Just to name a few downsides: LHEF is unnecessarily bloated, uses a plain text output, and lacks good centralised documentation, where even the existing definitions are unclear and therefore not widely followed in the physics community.

structured in different data sets, which are specially designed for quick read and write access even when using multi-threading[5], which is particularly important for parallelised high-performance computing. To run HEJ on computing grids a backend to pyHepGrid [124] was developed, which provides a common interface to run grid jobs on both ARC [125] and DIRAC [126]

Beside the LHEF based formats, HEJ can write out the generated events in the HepMC format version 2 [127] or 3 [128]. Alternatively events can also be directly analysed—without explicitly being written to disk—either with a custom analysis or through Rivet [129, 130]. Internally HEJ uses the `FastJet` library [69] for jet clustering, and LHAPDF [61] for PDFs. As pseudo-random number generators HEJ can use MIXMAX [131, 132] or RANLUX64 [133] through an interface to CLHEP [134, 135]. All optional interfaces can be selected at runtime through a YAML [136] configuration file, which is read via the yaml-cpp parser [137]. A detailed guide for all settings is given in ref. 5.

HEJ 2 can not only be used as a standalone executable, but also as a C++ library. Amongst others, this library is required for a reimplementation of the HEJ and Pythia merging of ref. 99 for arbitrary processes, which is currently under development. Furthermore, one can directly access the (Born-level) HEJ matrix elements to build a fixed-order Monte Carlo generator, called HEJ FOG. When using HEJ FOG as an input to HEJ—as the Born event in eq. (3.1.4)—one gets the pure, unmatched HEJ cross section. Hence, HEJ FOG is intended as an alternative for LO calculations in the merging, when these become prohibitively slow. For example in a setup similar to ref. 121, generating $10^6$ (partially unweighted) events for $W^+$ bosons with 9 jets takes $\mathcal{O}(15\,\text{min})$ with HEJ FOG on a single thread of a 2.4 GHz laptop CPU, while ref. 121 reported a total of $\mathcal{O}\left(10^5\,\text{h}\right)$ for the LO result with Comix [138] on a computer cluster.[6] Spending that much time on a process with only a small impact on the final cross-section is unnecessarily wasteful; a rough estimate is sufficient.

To further improve the speed of HEJ and to avoid code duplication, we are currently rewriting the handwritten implementation of the HEJ matrix elements in the symbolic algebra language Form [139, 140]. In Form one can write the currents at amplitude level, close the to the form reported throughout this thesis. The contraction and symbolic manipulation of these currents—by using the relations from appendix A—are done automatically at compile time. The resulting expressions for the matrix elements are then exported to C++ code, simplifying the runtime calcula-

[5] Currently HEJ is not multi-threaded, thus it does not utilise the full potential of the HDF5 format. But since the bottleneck when running HEJ is typically the fixed-order calculation, having a file format better suited for fixed-order is still desirable.

[6] Clearly this is an unfair comparison, since it says nothing about the integration performance. It should just be seen as illustration of "how much" simpler the HEJ matrix elements are.

tions down to (a minimal number of) complex-number multiplications. In the future this procedure may be generalised to also produce sub-subleading contributions that can be built up from already known currents, such as $qQ \to gqQg$ which contains two unordered currents.

HEJ calculates the cross section as the *sum of weights*, which differs from eq. (2.4.1) where we assumed the *average weight* to be the cross section—obviously both descriptions are interchangeable when knowing the number of weights. The advantage of using the sum of weights is that we do not have to write out zero-weight events when they fail the selection criteria. Formally the HEJ cross section is

$$\sigma_N = \sum_{i=1}^{N} x_i \sum_{j=1}^{M} y_{i,j} = \sum_{i=1}^{N} \sum_{j=1}^{M} w_{i,j} \,, \tag{3.1.14}$$

where $x_i$ are the weights as produced from the fixed-order generator, and $y_{i,j}$ the reweighting factors from HEJ. The full weights written out after resummation are then $w_{i,j}$. Here we take $N$ fixed-order events as an input and adding $M$ trail emissions during the resummation. We can estimate the Monte Carlo variance to be

$$s^2_{N,M} = M^2 \sum_{i=1}^{N} (x_i - \bar{x})^2 \, \bar{y_i}^2 + \sum_{i=1}^{N} x_i^2 \sum_{j=1}^{M} \left(y_{i,j} - \bar{y}_i\right)^2 \tag{3.1.15}$$

$$\leq \sum_{i=1}^{N} \left( \sum_{j=1}^{M} w_{i,j} \right)^2 + \sum_{i=1}^{N} \sum_{j=1}^{M} \left(w_{i,j}\right)^2 , \tag{3.1.16}$$

where $\bar{x}$ and $\bar{y}_i$ are the average values of $x_i$ and $y_{i,j}$ respectively. The second line gives a strict overestimator, which only depends on the full event weight after resummation. It is therefore not necessary to explicitly write out $x_i$ or $y_{i,j}$.

Importantly eq. (3.1.15) is always bigger then the uncertainty from just fixed order or just reweighting alone. Still, according to the first term, the total variation always scales like $\mathcal{O}\left(1/\sqrt{N}\right)$. In the limit $M \to \infty$ the total standard deviation is that of the fixed order input. For $M = 1$ we exactly reconstruct the old matching where one had one resummation point per fixed-order event. This is the worst-case scenario with the maximal uncertainty, since the second term would be at its maximal size. For any other $M > 1$ the second term decreases with $(NM)^{-1}$, leading to a smaller overall variation. This is equivalent to other Monte Carlo methods, which do not affect the fundamental scaling of the convergence with $N$, but give a smaller pre-factor.

Thus we have to choose an optimal $M$ which balances the time spent between integrating LO and resummation. In fig. 3.6 we find this value to be at around $M = 100$ trial emission per fixed order event.[7] At lower values one sees artefacts

[7] In principle one could choose different $M$ for different fixed-order events, to further optimise the integration

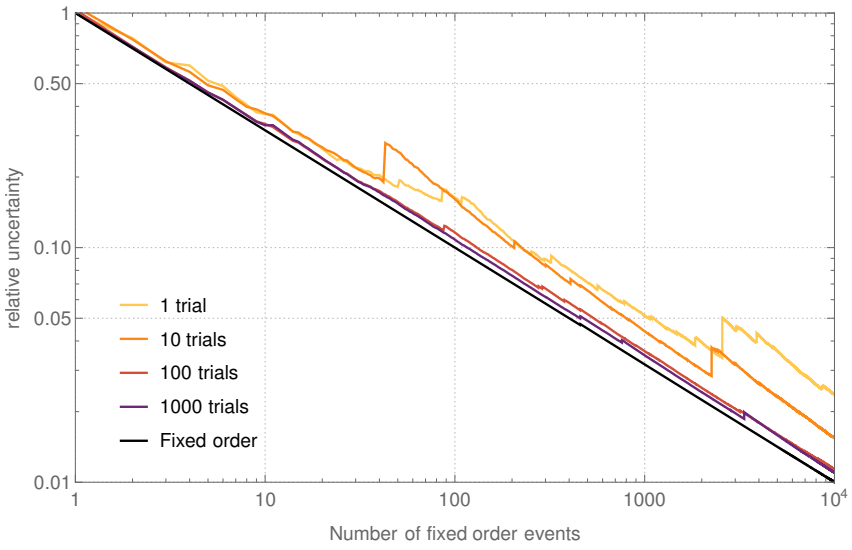



Figure 3.6: Scaling of the estimated standard derivation with increasing number of fixed-order events for $H + jj$ production. As the number of trials increases (from yellow to violet) one gets closer to the fixed order uncertainty (black line).

from hitting phase space region that is difficult to sample in the resummation (jumps in fig. 3.6). Higher values of $M$ show diminishing returns; with $M = 100$ we spend roughly 6% of the total CPU time on generating the resummation for $H + 2j$, thus increasing $M$ to 1000 leads to $\sim 50\%$ increased total runtime.

## 3.2 Results

In this section we discuss the results obtained with the new formalism in the key process of Higgs-boson production in association with at least two jets. Firstly we confirm that if we limit ourselves to matching with fixed-order samples with up to three jets that we reproduce the results obtained with the previous formalism, but now with a much higher efficiency. We will then show and discuss the impact of being able to increase the multiplicity in the fixed-order samples and also compare our results to NLO predictions.

The matching procedure described in this work is significantly more efficient and flexible than the approach used in previous versions of HEJ. To illustrate this, we present new results for the production of a Higgs boson in association with at least two jets matched to LO events with up to four jets. Previously, matching of HEJ to just three jets was achieved for this process, while using significantly more CPU resources than necessary with the current approach. In its new formulation,

the matching is in practice only limited by the capabilities of the underlying fixed-order generator. For instance, the generation of one set of 1000 unweighted LO events for the production of a Higgs boson with four jets typically took a few CPU days using `MadGraph5_aMC@NLO` [55]. It is then just a few additional CPU seconds to generate 100 weighted resummation events from each of the fixed-order 4-jet events, so 100 000 weighted trial resummation configurations in total. In the previous matching approach, generating 100 000 resummation configurations would require the same number of computationally expensive fixed-order matrix element evaluations.[8] Since the resummation is not followed by any computationally intensive steps, we only consider the generation of weighted events here. Nonetheless, we also observe a marked improvement in a short test simulation with unweighted events.

This section will present the results obtained with the new procedure for matching and resummation. Section 3.2.1 describes the cuts and analysis used. Section 3.2.2 compares new results with matching up to three jets with those obtained previously, and demonstrates that the two methods yield equivalent results. Section 3.2.3 investigates the stability of the results obtained by investigating the impact of increasing the order to which matching is achieved. In general, the matching to higher multiplicities should have little impact for configurations where the four-jet contribution is insignificant or the approximation within HEJ already provides a good description. Conversely, the corrections from matching to successive multiplicities can serve to indicate the stability of the HEJ predictions for observables sensitive to additional hard radiation. Finally, in section 3.2.4 we match the inclusive Higgs-plus-dijet cross section to NLO accuracy, thus obtaining the most precise predictions for $H + jj$-production, including the effects of VBF cuts. These results are compared to those obtained at fixed next-to-leading order accuracy.

### 3.2.1 Setup

To facilitate the comparison with previous results we will adopt the cuts of the ATLAS experimental analysis of ref. 141, and the parameters of our analysis in ref. 98. The analysis was implemented in Rivet version 2 [129]. To recapitulate, we consider the gluon-fusion-induced production of a Higgs boson together with at least two anti-$k_t$ jets with transverse momenta $p_{j\perp} > 30\,\text{GeV}$, rapidities $|y_j| < 4.4$, and radii $R = 0.4$ at the 13 TeV LHC. While it is unrelated for the considerations of the QCD corrections considered in this chapter, we consider the Higgs-boson decay

[8] Note that for our specific setup a direct comparison between both approaches is not possible since there is no analogue to the concept of unweighted fixed-order events in the "old" approach.

into two photons with

$$\begin{aligned} |y_\gamma| < 2.37, \qquad & 105\,\text{GeV} < m_{\gamma_1\gamma_2} < 160\,\text{GeV}, \\ p_{\gamma_1\perp} > 0.35\, m_{\gamma\gamma}\,, \qquad & p_{\gamma_2\perp} > 0.25\, m_{\gamma\gamma}\,, \end{aligned} \tag{3.2.1}$$

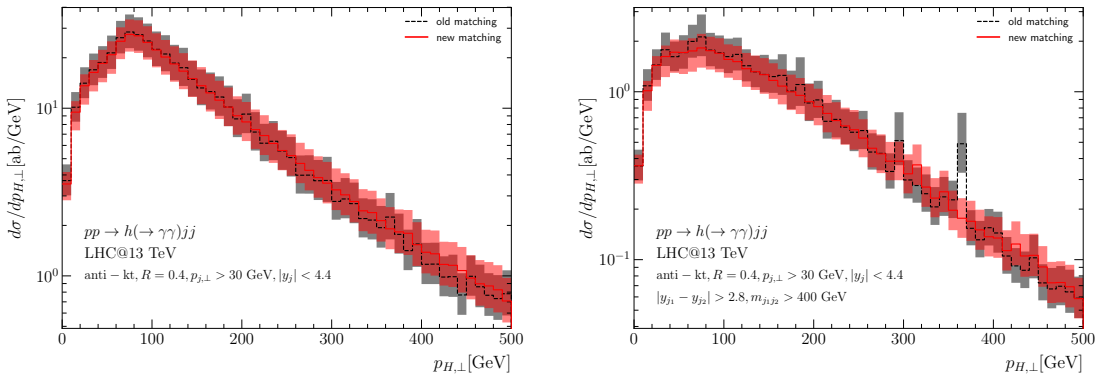
and separations $\Delta R(\gamma, j), \Delta R(\gamma_1, \gamma_2) > 0.4$ from the jets and each other. To be consistent with our previous analysis we set the Higgs-boson mass to $m_H = 125\,\text{GeV}$, a width of $\Gamma_H = 4.165\,\text{MeV}$ and a branching fraction of 0.236% for the decay into two photons. We use the CT14nlo PDF set [142] as provided by LHAPDF version 6 [61]. In addition to inclusive quantities with the basic cuts listed above, we also consider additional VBF-selection cuts applied to the hardest jets as in ref. 141:

$$|y_{j_1} - y_{j_2}| > 2.8, \qquad m_{12} > 400\,\text{GeV}. \tag{3.2.2}$$

In the first step, we generate LO events with two, three and four jets. With our new matching procedure we are free to use an arbitrary fixed-order event generator for this purpose. For the present analysis we employ version 2.5.5 of `MadGraph5_aMC@NLO` [55]. For each jet multiplicity we produce about 2000 sets of unweighted events, each comprising 10 000 events for the sets with two or three jets and 1000 events for sets with four jets.

As the transverse momenta of the jets are modified during resummation through eq. (3.1.6), we have to generate at least a fraction of events with Born-jet momenta below the threshold of 30 GeV required from the observed jets. As already shown in fig. 3.1 the contribution after the resummation from such tree-level configurations in the matching drops off very rapidly below the jet transverse momentum analysis scale of 30 GeV. Passing this information to the underlying fixed-order generator, such that only a small fraction of events are generated below the nominal transverse momentum threshold could improve the sampling efficiency considerably. Having such an option would therefore be highly desirable. For the time being, we manually generate 200 additional sets of Born-level events with transverse momenta down to 20 GeV for each jet multiplicity.

Events with more exclusive jets than can be reasonably evaluated at leading order in `MadGraph5_aMC@NLO` are unmatched and generated with HEJ FOG instead of full leading order. In this way, we can supplement the fixed-order input with events including up to ten jets obtained within the HEJ approximation. These events are simply passed through the same matching mechanism based on eq. (3.1.4) just as the lower-multiplicity events obtained using `MadGraph5_aMC@NLO`. The maximum multiplicity of ten is an arbitrary cut-off, based on an explicit check that the impact on the shown observables at this multiplicity is negligible.




(a) (b)

Figure 3.7: Comparison of the new matching procedure to previous HEJ results obtained in ref. 98. The panels show the transverse momentum distributions of the Higgs boson for (a) inclusive cuts and (b) VBF cuts.

Since the final kinematics required for a kinematic scale setting are not known at the point of generating fixed-order events, we use a fixed renormalisation and factorisation scale of $\mu_r = \mu_f = m_H$ during the fixed-order generation. After resummation the events are rescaled to a central scale of $\mu_r = \mu_f = H_T/2$. In order to assess the scale dependence, we independently vary both the renormalisation and factorisation scales by factors $\{1/2, 1/\sqrt{2}, 1, \sqrt{2}, 2\}$ and discard combinations with $\mu_r/\mu_f < 1/2$ or $\mu_r/\mu_f > 2$. In the effective Higgs-gluon coupling, we keep the renormalisation scale at the Higgs-boson mass and apply the limit of an infinite top-quark mass. These scale settings and even the use of the infinite top-mass limit are however not inherent to the use of the high-energy resummation of HEJ, but can be included with modified components of the amplitudes, similar to ref. 109.

After generating the tree-level input events, we apply resummation as presented in the previous sections. Recent progress described in ref. 98 allows us to apply resummation not just for FKL-ordered matching-events, but also the sub-leading contribution from events with three jets or more, where the rapidity-ordering of the two most forward or most backward jets is flipped compared to FKL ordering. This corresponds to a gluon emission outside a rapidity-interval delimited by quark jets. For each resummation-type tree-level event, we generate 100 weighted trial configurations in the resummation phase space. For the remaining sub-leading events we cannot add resummation and simply adjust the factorisation and renormalisation scales as described above.

### 3.2.2 Comparison to previous results

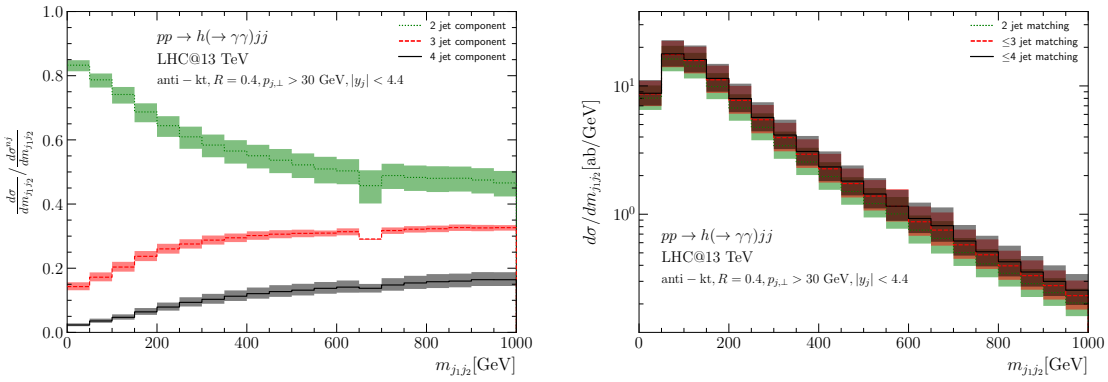
In order to demonstrate the validity of the new approach we compare here first our results with LO matching up to three jets to those obtained in our previous work ref. 98 . We find good agreement within the statistical errors. As examples, we show the transverse momentum distributions for the Higgs boson for inclusive and VBF cuts in fig. 3.7. The previous and new method of organising the calculation are equivalent. For the comparison, we have adjusted our settings to match those in ref. 98 as closely as possible. Apart from restricting the fixed-order matching to configurations with at most three jets, this also means that the extremal partons are required to have a fixed minimum transverse momentum of 27 GeV instead of a fraction of the corresponding jet momentum, as discussed in section 3.1.

### 3.2.3 Impact of four-jet matching on distributions

The HEJ approximation is exact in the limit of multi-Regge kinematics, i.e. for large rapidity separations between hard jets. An equivalent characterisation is to demand the centre-of-mass energy and the invariant masses between all final-state jets to be much larger than the typical transverse momenta of these. If these conditions are fulfilled, we expect HEJ to produce accurate predictions and hence small matching corrections. In order to assess the perturbative stability of the final predictions, we will here study the impact on the resummed and matched cross section of scale variations and of successive matching to two-jet, three-jet and four-jet tree-level events.

One of the main goals of HEJ is to improve the prediction of the gluon-fusion background to Higgs-boson production in weak-boson fusion. Standard VBF cuts project out a kinematic region with a large invariant mass between the hardest jets, where the gluon fusion receives significant contributions from higher jet multiplicities. Figure 3.8a displays the relative contribution of the exclusive two-, three- and four-jet component to the distribution on the invariant mass between the two hardest (in transverse momentum) jets. The relative contribution from exclusive three- and four-jet-events increases with increasing $m_{12}$. Figure 3.8b displays the impact of matching of successive multiplicity on the distribution of the invariant mass between the two hardest (in transverse momentum) jets. The effect of the four-jet matching is small but non-zero even at large $m_{12}$. This is because even in this limit a large separation between *all* jets is not guaranteed.

The contribution from jet multiplicities of more than or equal to 5 is less than 5% for an invariant mass of at least 1 TeV. We conclude that the uncertainty on the




(a) (b)

Figure 3.8: Distribution of the invariant mass between the hardest jets. Panel (a) shows the fractional contributions from exclusive two-, three-, and four-jet events. Panel (b) depicts the effects of fixed-order matching up to two, three, and four jets.

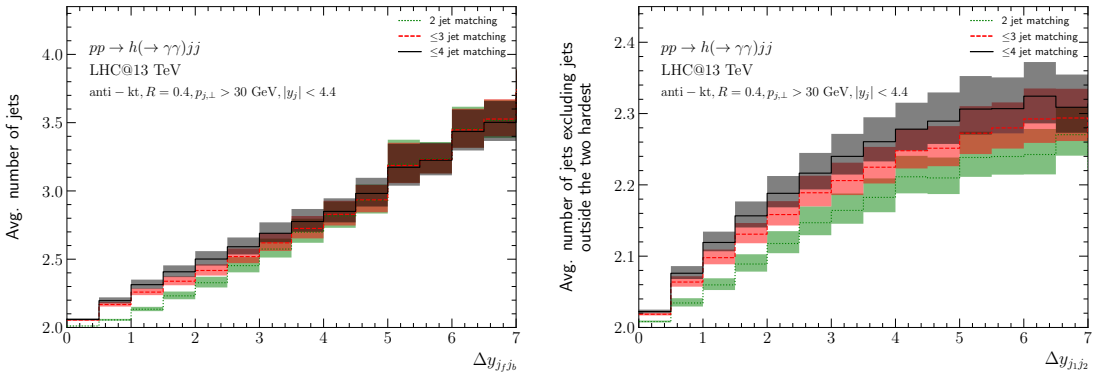
distribution of $m_{12}$ from terminating the matching at the four-jet contribution is insignificant for the study here, and well within the quoted scale variation.

A central prediction of BFKL, which arises also within HEJ, is a linear increase in the number of jets for a growing rapidity span between the most backward and forward jets.[9] This behaviour is demonstrated in fig. 3.9, which also investigates the impact of matching to tree-level of successive multiplicities. Although the contribution from higher jet multiplicities increases with the rapidity separation, the effect of fixed-order matching on this observable actually decreases. This confirms our expectation that the HEJ approximation works well for large $\Delta y_{j_f, j_b}$. It is this linear increase in the average number of jets versus increasing rapidity span which can be exploited to suppress the gluon-fusion contribution with a central jet veto [85].

In contrast to this, if the two *hardest* jets are tagged, and only jets in-between these are counted as a function of the rapidity difference between the hardest jets, then the initial linear growth stalls at an average number of jets of around 2.3. The difference in behaviour to the VBF contribution is therefore less pronounced by tagging the hardest jets, rather than the most forward and backward hard jets. This was investigated further in ref. 85. Also, the impact of the matching corrections remains sizeable for all rapidity separations.

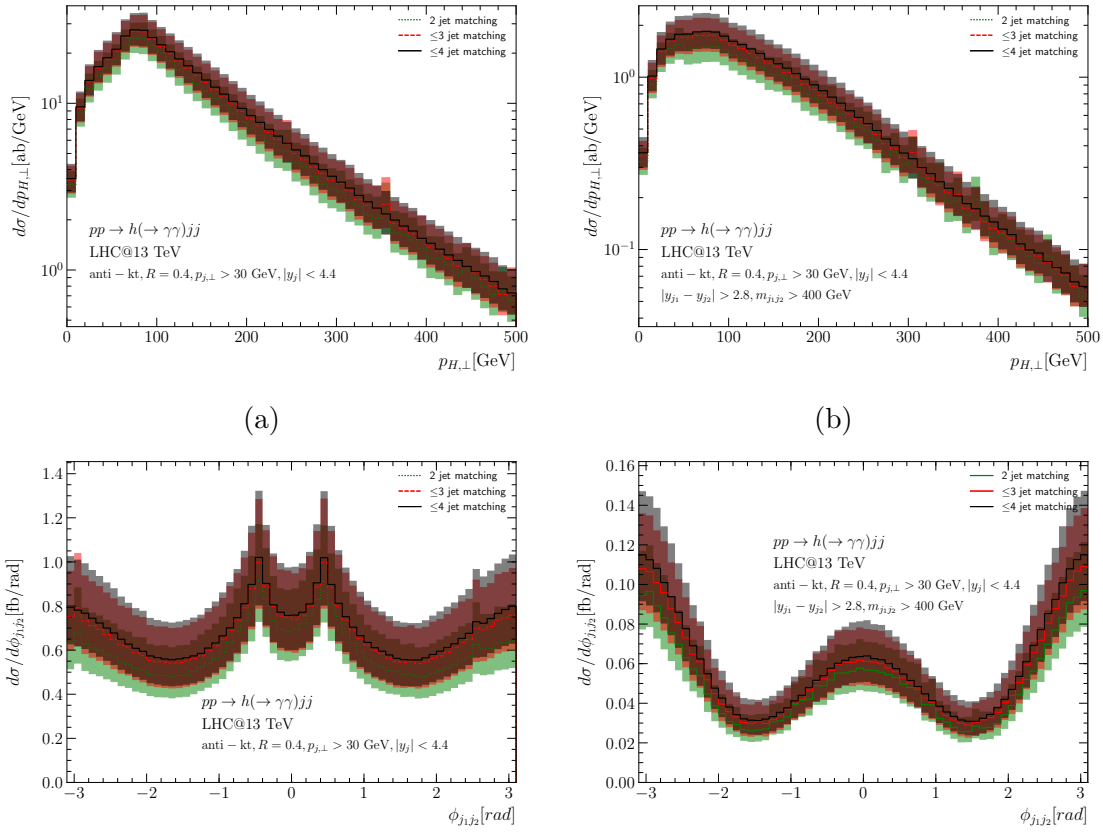
In observables which are neither dominated by higher jet multiplicities nor completely

[9] This growth continues until the invariant mass of just the forward and backward jets is so large that no other jets can be emitted due to energy and momentum constraints. The fixed-order NLO results have a similar behaviour at small $\Delta y_{j_f, j_b}$ until the average jet multiplicity is saturated by the fixed-order truncation.




(a) (b)

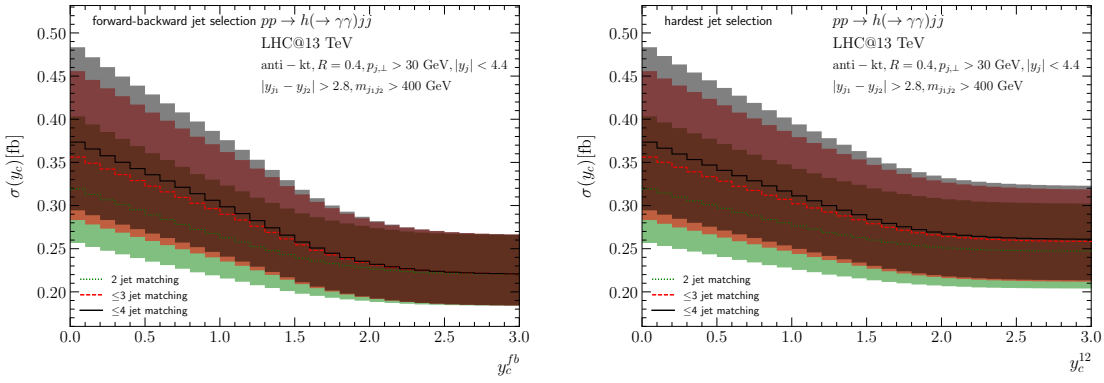
Figure 3.9: Average number of jets for fixed-order matching up to two, three, and four jets. In (a) we show the average total number of jets over the maximum rapidity-separation. In (b) we show the number of jets in the rapidity region of the two hardest jets.




(a) (b)

(c) (d)

Figure 3.10: Distribution of the Higgs-boson transverse momentum and the azimuthal angle between the two hardest jets (top and bottom respectively). The results on the left are with inclusive cuts, while on the right we choose the VBF selection.




(a) (b)

Figure 3.11: Effect of a jet veto between (a) the most forward and backward jets and (b) the two hardest jets. Events with additional jets within a distance of $y_c$ to the rapidity centre are discarded.

described by the HEJ approximation we observe that the matching corrections are converging, but the corrections from four-jet matching are non-negligible. In first two plots of fig. 3.10 we show the distribution of the Higgs-boson transverse momentum with inclusive and with VBF cuts. While there is a notable difference between the matching to fixed-order predictions up to two and three jets, the effect of four-jet matching is much smaller. In all cases the matching corrections are well inside the scale variation.

The azimuthal angle between jets is of particular interest for the extraction of the $CP$-properties of the effective coupling between the Higgs boson and gluons. The second row of fig. 3.10 shows the effects of fixed-order matching on the distribution of the angle between the two hardest jets. Similar to the transverse momentum distribution, the corrections from four-jet matching are uniformly moderate.

In order to achieve a greater reduction of the gluon-fusion background to weak-boson fusion within the fiducial region, a veto on further jets can be applied. This has the added benefit of reducing the contribution from higher jet multiplicities, which is more challenging to predict in perturbation theory. The effectiveness of such a cut relies on the difference in the quantum corrections to the processes of VBF and GF [114]. Since this difference is due to the $t$-channel colour-octet exchange of the GF process, we will apply a central jet veto only in the regions away from the tagging jets, since the collinear regions have similar emissions in VBF and GF. This is a slight improvement on the normal central jet veto cuts, and is inspired by the Zeppenfeld variable [143]. Here, we consider a veto of events with jets within a rapidity distance $y_c$ to the rapidity centre of either (a) most forward and backward jets or (b) the hardest jets (see also ref. 85, 143 and section 4.3.4). In case (b), we

only consider vetoing on further jets which are in between the two hardest jets. The results are shown in fig. 3.11. As expected, the cross section in case (a) converges for large $y_c$ to the exclusive prediction for the production of a Higgs boson with exactly two jets, irrespective of the fixed-order matching to higher multiplicities. When applying the jet veto between the hardest jets in case (b), the overall reduction in the GF component is considerably smaller.

### 3.2.4 Matching and comparison to next-to-leading order

The complete reformulation of the formalism for matching and all-order summation described in section 3.1 has allowed for matching to higher jet multiplicities in HEJ. The impact of the four-jet matching on the studied distributions is small. The method presented in the earlier sections has been concerned with a point-by-point matching of the resummation to full high-multiplicity tree-level accuracy. As extensively demonstrated in section 3.2.3, this achieves perturbatively stable results for the shapes of distributions. In order to reduce the scale variation and benefit from full NLO results for $H + jj$-production, we will now rescale the results for HEJ within the inclusive cuts of eq. (3.2.1) to the NLO cross section for each choice of factorisation and renormalisation scale. Thereby, full NLO accuracy is obtained for all dijet observables, LO accuracy for trijet observables, and the impact on the shape of distributions from four-jet contributions is accounted for at LO. This method was applied also in ref. 85. While this approach does not change the shape of distributions, the scale variation is reduced to the level of NLO predictions. We will here compare these predictions to those obtained at fixed NLO using MCFM [54, 84] and Sherpa [29, 92] together with OpenLoops [53, 144].

Figure 3.12 compares the predictions for the distribution of the invariant mass between the two hardest jets. The scale variation on the HEJ results is vastly reduced to that of fig. 3.8, as generally expected by the inclusion of the full NLO corrections. The distribution obtained with HEJ for the invariant mass between the two hardest (in transverse momentum) jets is still significantly steeper than that at pure NLO, as a result of the inclusion of significantly higher jet multiplicity, and the fact that hard central jets have a slightly smaller PDF-suppression than hard forward jets, and therefore the two hardest jets tend to also be central. This means that the predictions for the cross section within the VBF cuts is significantly smaller with HEJ than for NLO, and indeed lies outside the scale-variation band obtained at NLO.

In numbers, the cross sections obtained (at NLO) for $pp \to H(\to \gamma\gamma)jj$ for inclusive cuts and with a central scale choice of $\mu_r = \mu_f = H_T/2$ is $6.58^{+0.08}_{-0.57}$ fb. This is obviously

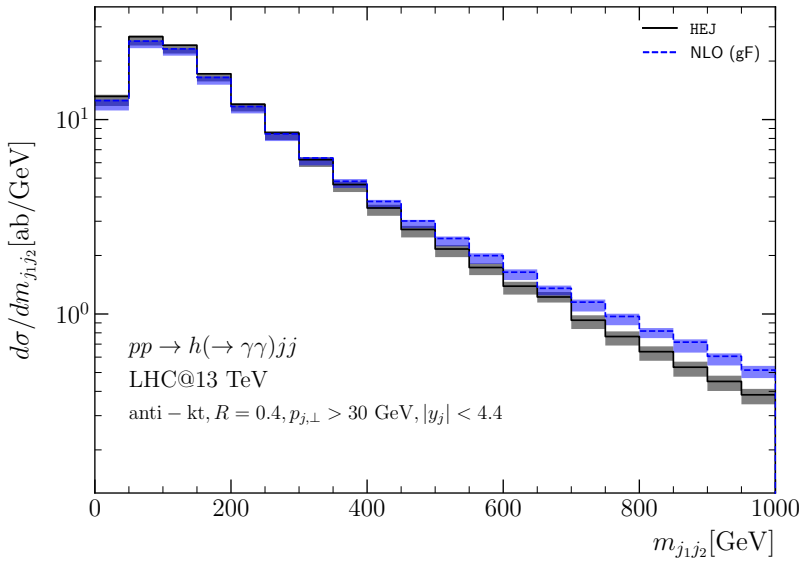



Figure 3.12: Distribution of the invariant mass between the two hardest jets. The HEJ result with fixed-order matching up to 4 jets is shown by the black, solid line. The MCFM NLO prediction corresponds to the blue, dashed line. The central renormalisation and factorisation scale is set to $H_T/2$, and they are varied independently by a factor of two.

the same as that obtained with HEJ for inclusive cuts, once the cross sections are normalised to NLO accuracy. For the VBF cuts, the NLO cross section is $0.87^{+0.02}_{-0.09}$ fb, and that obtained for HEJ is $0.56^{+0.03}_{-0.07}$ fb. Even though the inclusive cross section for HEJ is normalised to that obtained at NLO, a sizeable difference in the cross section within the VBF cuts arises due to a difference in the slope of distribution in $m_{12}$ and the requirement of $m_{12} > 400\,\text{GeV}$ for the VBF cuts. The VBF cuts cause a similar reduction in the cross section to 13.2% (NLO) and 8.5% (HEJ) of the inclusive cross section respectively.

Comparing the results of figs. 3.12 and 3.13 we observe that a choice of a central scale for the NLO calculation of $\mu_r = \mu_f = H_T/2$ leads to a suspiciously small scale variation—and indeed the central scale choice gives results close to the extremum obtained with the variations, despite the scales being varied either side of the central choice of $\mu_r = \mu_f = H_T/2$. Such a behaviour of the scale variation often indicates that the NLO scale variation obtained with this scale choice is underestimating the theoretical uncertainty [145].

In fact, ref. 145 investigated the distribution in $m_{12}$ for dijet production at NNLO at the LHC, and found that at large $m_{12}$ this scale choice is favoured over scales based on jet $p_\perp$ for perturbative convergence. The invariant mass between the two

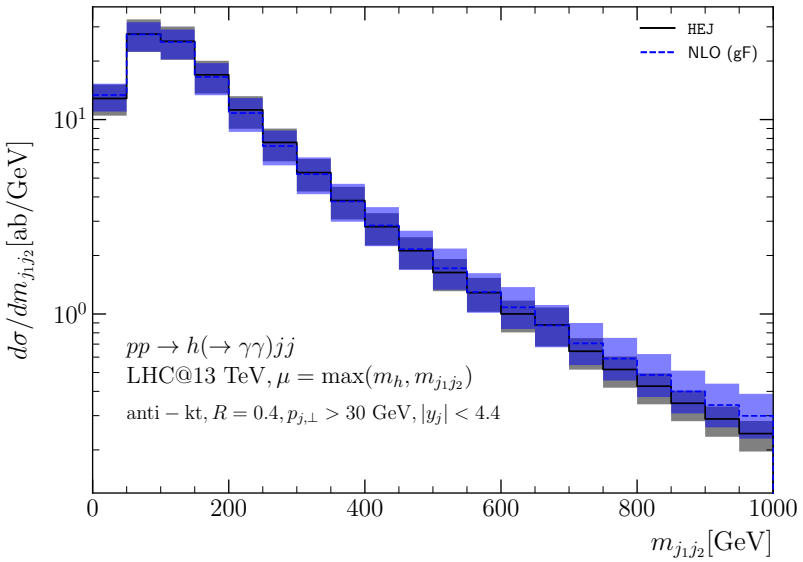



Figure 3.13: Distribution of the invariant mass between the two hardest jets, for a central scale choice of $m_{12}$. The HEJ result with fixed-order matching up to 4 jets is shown by the black, solid line. The SHERPA NLO prediction corresponds to the blue, dashed line.

hardest jets obviously is not a stable perturbative scale choice for all bins in the distribution, which extends to very low values of $m_{12}$. With a central scale choice of $\mu_r = \mu_f = \max(m_H, m_{12})$, the central scale choice leads to predictions in the centre of the variation band. The scale variation bands obtained with NLO and HEJ also overlap in each bin of the distribution. With this central scale choice, the cross sections obtained at NLO for inclusive cuts is $6.23^{+1.11}_{-1.22}$ fb. For the VBF cuts, the NLO cross section is $0.54^{+0.16}_{-0.12}$ fb, and that obtained for HEJ is $0.36^{+0.05}_{-0.06}$ fb. The VBF cuts cause a similar reduction in the cross section to 8.7% (NLO) and 5.8% (HEJ) of the inclusive cross section respectively.

It is worth noting that, since at Born-level $m_{12}$ is given by eq. (2.1.5), a central scale choice of $\mu_r = m_{12}$ systematically runs $\alpha_s$, such that $\alpha_s \Delta y_{j_1j_2}$ tends to a constant for large $\Delta y_{j_1j_2}$. This would seem to spoil the standard argument of BFKL noting large and systematic leading logarithmic corrections of the form $(\alpha_s \Delta y_{j_fj_b})^k$ at large $\Delta y_{j_fj_b}$. At least for $\Delta y_{j_1j_2}$ sufficiently large, $m_{12}$ is close to the hadronic collision energy that only two jets exists, because hard radiation, beyond the two required jets, is suppressed. For events with more than two jets, there is no direct correlation between $\Delta y_{j_1j_2}$ and $\Delta y_{j_fj_b}$. The results for the scale choice of $\mu_r = \mu_f = \max(m_H, m_{12})$ are discussed further in the next chapter. Here we just note that the apparent convergence of the perturbative series—i.e. a comparison of the LO and NLO results and scale variation—is not significantly different for the two scale choices. The scale

variation around $\mu_r = \mu_f = H_T/2$ is accidentally small, since the central scale choice leads to the maximum cross section within the variation.

Figures 3.14a and 3.14b investigates the potential for using perturbative corrections in the form of additional jet-radiation as a means of identifying the gluon-fusion production channel. For the same event selection, the figure compares the results for the average number of jets counting additional jets (a) between the most forward and backward jets, and (b) in-between the two hardest jets only. The results of fig. 3.14a are relevant for e.g. jet vetos between the most forward and most backward hard jet, whereas fig. 3.14b is relevant if the veto is applied between just the two hardest jets in the event. The results for HEJ are identical to those for 4-jet matching in fig. 3.9 (since just the total cross section has been adjusted to the NLO result for $H+jj$), but the results are here compared to those obtained using the NLO calculation for QCD $H + jj$-production. For reference we also show NLO results for the VBF channel.

Over the full range the VBF channel shows only mild increase in the number of jets, with only a small scale uncertainty. In contrast, as observed also in previous analyses [146], the results obtained for gluon-fusion at NLO tends towards 2.5, where the exclusive, hard three-jet cross section is as large as the two-jet cross section. This clearly illustrates the slow convergence of the perturbative series for gluon-fusion and the quick convergence for VBF. The results for NLO and for HEJ begin diverging already at small $\Delta y_{j_f j_b}$ above $\Delta y_{j_f j_b} > 1$. It is worth noting that the linear growth in the number of hard jets over $\Delta y_{j_f j_b}$ has been experimentally confirmed for several processes with colour octet exchanges in the $t$-channel [91, 94].

Even though exactly the same events are involved, the breakdown of the convergence is less obvious in fig. 3.14b. The number of jets in-between the two hardest is obviously smaller, and both the results for HEJ and for NLO appear to asymptote to a value for the average number of jets of 2.2 for NLO and 2.3 for HEJ. In both selections, hardest or most forward and backwards, any cut on the third jet would predominantly affect the gluon fusion-channel, while keeping most of the VBF cross section.

Figures 3.14c and 3.14d also show the predictions for the Higgs transverse momentum spectrum obtained at NLO and with HEJ both for inclusive and VBF-cuts. The distributions are very similar for inclusive cuts, with a peak around 80 GeV, and the spectrum from HEJ slightly harder. For VBF cuts, the prediction for HEJ is lower than that for NLO, as a result of the steeper spectrum in $m_{12}$ and the requirement of $m_{12} > 400\,\text{GeV}$. The two predictions for the high-$p_\perp$ tail within the VBF cuts coincide, but in this region the infinite top-mass approximation is certainly not trustworthy.

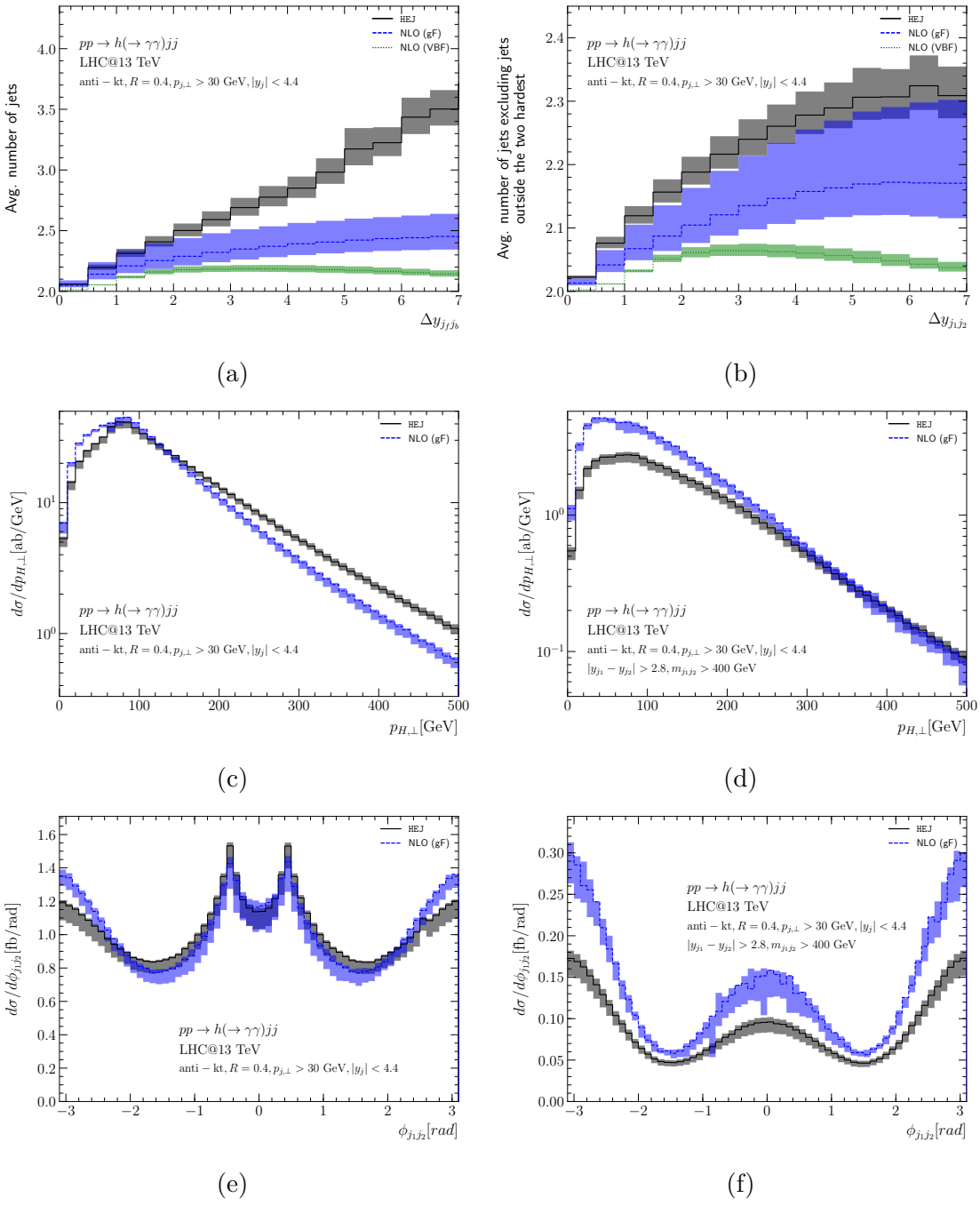



Figure 3.14: Comparison of HEJ results with fixed-order matching up to 4 jets (black, solid) with NLO predictions from MCFM for gluon fusion (blue, dashed) and weak-boson fusion (green, dotted). The shown observables are (a) the average jet multiplicity, (b) the number of jets in between the two hardest jets, the distribution of the Higgs-boson transverse momentum with (c) inclusive and (d) VBF cuts, and the distribution of the azimuthal angle between the hardest jets with (e) inclusive and (f) VBF cuts.

Finally, figs. 3.14e and 3.14f compare the azimuthal angle between the two hardest jets again for inclusive and VBF cuts respectively. In both the distributions, the region of back-to-back jets at $\phi = \pm\pi$ is slightly suppressed in HEJ compared to NLO. The valley at $\phi = 0$ for the inclusive cut is due to the jet-algorithm removing the collinear region. The result within the VBF cuts on fig. 3.14f show that the reduction in the cross section within the VBF cuts for HEJ compared to NLO predominantly is in the region of back-to-back jets and jets in the same azimuthal direction. The region at $\phi_{j_1 j_2} = 0$ is not collinear within the VBF cuts, and so the structure induced by the jet algorithm within the inclusive cuts of fig. 3.14e is not present within the VBF cuts of fig. 3.14f.

## 3.3 Summary

We have presented a reformulation of the matching formalism within HEJ, which recasts the calculation as one of merging fixed-order samples of increasing multiplicity. The merging is performed respecting the resummation of perturbative terms logarithmically enhanced at large $\hat{s}/p_\perp^2$. While the formalism is mathematically equivalent to that previously used, stable results are obtained using orders of magnitude less CPU time. This allows matching to be performed to higher multiplicity.

The new formalism has been used in a study of Higgs-boson production in association with dijets. The impact of the higher-multiplicity merging is minimal on the shape of distributions important for the application of VBF cuts. For a central scale choice of $\mu_r = \mu_f = H_T/2$, the VBF cuts reduce the inclusive cross section of $6.58^{+0.08}_{-0.57}$ fb on the $H \to \gamma\gamma$-channel to 13% ($0.87^{+0.02}_{-0.09}$ fb) at NLO, or 8.5% ($0.56^{+0.03}_{-0.07}$ fb) once both NLO and the HEJ-corrections are accounted for. The further suppression within HEJ is due to a more steeply falling invariant mass spectrum between the two hardest jets. The NLO scale dependence is estimated by varying the renormalisation and factorisation scale independently by a factor of up to two. However, the scale variation around $H_T/2$ is artificially small, since the central scale choice achieves a value close to the maximum within the variations. With a scale choice of $\mu_r = \mu_f = m_{12}$ (but bounded from below by $m_h$), the spectrum is similar at NLO and with the further HEJ-corrections. With the scale choice, the inclusive NLO cross section for $H(\to \gamma\gamma)jj$-cross section is $6.23^{+1.11}_{-1.22}$ fb, and $0.54^{+0.16}_{-0.12}$ fb within the VBF cuts. The result for HEJ within the VBF cuts is $0.36^{+0.05}_{-0.06}$ fb. The VBF cuts cause a similar reduction in the cross section to 8.7% (NLO) and 5.8% (HEJ) of the inclusive cross section respectively.

The formalism presented here has been publicly released in version 2 of HEJ [3], and will be used as the basis for all expansions to the HEJ. As an example the heavy-

quark mass effects presented in the next chapter have been implemented within the new framework.

# Chapter 4

# Finite quark mass effects in HEJ

In this chapter we calculate the gluon-fusion initiated contribution of Higgs-boson production in association with two jets supplementing the fixed-order results with the leading logarithmic corrections in $\hat{s}/p_\perp^2$ to all orders in the coupling *and* including full quark mass effects. The necessary mass-dependent components for the resummation are derived, allowing us to perform a first calculation of the interference of top- and bottom-quark mass contributions in $H + 2j$.

Let us first review the current status of the calculation of the quark-mass effects in Higgs-boson production. Higgs bosons are produced at the greatest rate at the LHC through heavy quark mediated gluon fusion, where the Born-level process is loop-induced and at order $\alpha_s^2$ in massless QCD. Since the coupling of the Higgs boson to a quark is proportional to the quark mass, the gluon-fusion process is dominated by the contribution from a top-quark loop. Inclusive production is known differentially to N$^3$LO (to order $\alpha_s^5$) in the limit of infinite top-mass [147–150]. Finite top-mass effects in the inclusive cross section can be taken into account at one order lower (to N$^2$LO in $\alpha_s$) by a formal expansion in $(m_h/m_t)$. The effects on the total cross section of the finite top-mass are found to be very small indeed [151–153]. An explicit calculation of the loop contribution using the full propagator dependence allows also the inclusion of the contribution from bottom-quarks, and it is found here that the bottom-top interference effects are of the order of $-5\%$ [149].

Higgs-boson production in association with *one* jet obviously forms a subset of the higher order radiative corrections to the calculation of inclusive Higgs-boson production. As such, it is known to N$^2$LO in $\alpha_s$ (to order $\alpha_s^5$) in the limit of infinite top-mass, and to NLO with full dependence on the heavy-quark propagator [154].[1]

[1] Reference 155 combined both calculation to get the 'current best prediction', by differentially reweighting N$^2$LO in the effective theory with the ratio of NLO at finite over NLO at infinite top mass. This is similar to the approach we will use in section 4.3 to produce improved NLO predictions for Higgs boson production with two jets.

This explicit calculation of the quark loops can be used to check the earlier reported top-bottom interference effects, which were approximated using a small-mass expansion for the amplitudes involving bottom-quarks, and the infinite mass limit for the top-quark contribution [156]. For $H + j$-production at NLO, the effect of the full dependence on the heavy-quark propagator momentum is a 9% *increase* in the overall cross section over the result obtained in the infinite top-mass limit. The approximation of the infinite top-mass limit in $H + j$-production is therefore missing a much larger contribution than for inclusive Higgs-boson production. Furthermore, there is a strong phase-space dependence: for transverse momenta of the Higgs boson larger than 800 GeV, the effects of the full dependence on the heavy-quark propagators leads to a *suppression* over the result for an infinite top-mass of more than an order of magnitude. For processes with more than just the Higgs boson in the final state, the limit of infinite top-mass loses not only the dependence on the mass of the propagating quarks, but also the full kinematic dependence on the propagators in the loop-diagrams.

Higgs-boson production in association with *two* jets can proceed through both of the processes of weak-boson fusion and gluon fusion. The VBF process is directly sensitive to the coupling between the Higgs boson and the weak bosons, whereas the GF $H + 2j$-process allow for studies of possible $CP$-admixtures in the Higgs sector [87, 157]. A precise study of either of these effects requires a separation of the contribution from the two processes, which has to be guided by detailed calculations. Luckily, these indicate that the interference between the two processes is negligible [81–83], so the two processes can in principle be studied independently.

The VBF process is known fully differentially at N$^2$LO [158, 159] (i.e. to order $\alpha_s^2\alpha^2$) and the inclusive cross section is known in the effective structure function approach to N$^3$LO [160] (i.e. to order $\alpha_s^3\alpha^2$). One important lesson from these calculations is that while the higher order perturbative corrections to the inclusive cross sections are very small indeed, within typical VBF-cuts the N$^2$LO-corrections to the NLO-result can be 3 to 4 times larger, and reduce the cross section by 4%, with effects of up to 7% on distributions.

The contribution through GF to $H + 2j$ is known with full dependence on the heavy-quark propagator just at LO [86, 161, 162], and at NLO in the infinite top-mass limit [54, 84]. The situation is the same for $H + 3j$ [163, 164]. In the current chapter we present a calculation of higher-order perturbative corrections to the GF component of $H+2j$-production, maintaining the full dependence on the heavy-quark propagator in the heavy-quark mediated coupling to the Higgs boson, and including the effects of propagating both top- and bottom-quarks. The results obtained are exact in the limit of large dijet invariant mass, which is relevant for the VBF- and

gluon-fusion CP-studies, and are furthermore matched to the highest-order fixed-order perturbative result which could be produced with available tools, in this case SHERPA [29, 92] in combination with OpenLoops [53, 144].[2] The results rely on the observation that the high-energy limit commutes with any limit taken on the masses of the propagating quarks in the coupling to the Higgs boson [109].

In section 4.1 we explore the structure of the amplitudes for the different subprocesses of $pp \to H + 2j$ with full dependence on finite quark masses. We use these results to construct matrix-elements within HEJ which contain all finite quark mass effects and maintain accuracy to leading logarithm in $|\hat{s}/\hat{t}|$ at all orders in $\alpha_s$. This manifestly includes the calculation of subprocesses with a high number of high-energy jets, going far beyond what is possible at fixed-order with finite quark mass effects. In section 4.2 we describe the different types of matching to fixed-order which we employ in the HEJ predictions. This is quite involved owing to the variety of fixed-order samples available. In section 4.3 we present our results, focussing separately on the effects of higher perturbative orders and the effects of finite top-mass, before we compare our most accurate HEJ prediction with the most accurate available fixed-order prediction. In section 4.4 we summarise our findings.

## 4.1 Quark mass effects in Higgs-boson production with HEJ

As discussed before in the high-energy limit amplitudes factorise into products of process-dependent currents and universal emission vertices. The production of a Higgs boson with at least two jets in HEJ has been described in the infinite quark mass limit in ref. 98. The only difference coming from the Higgs couplings were contained in the low multiplicity currents, without modifying the resummation. Thus, this $t$-channel factorisation of the amplitude lends itself to the inclusion of finite quark mass effects, because the diagrams and loops do not become any more complicated than those at leading order for arbitrarily many emissions. Finite quark mass effects in $H + 2j$ were studied in the high-energy limit in ref. 109. In this section, we recap some results and describe the necessary adaptation to incorporate the results in the current-structure of HEJ. We will start our discussion with the simplest LO configurations, before moving on to more complex arrangements. The generalisation to all order in $\alpha_s$, and thus the resummation of $\log|\hat{s}/\hat{t}|$ is then equivalent to section 2.3.

[2] This gives $pp \to H + 2j-$processes with full quark mass dependence. The corresponding results for $pp \to H + 3j$ in ref. 164 could have been included directly if the implementation was readily available or if the results were available as event files.

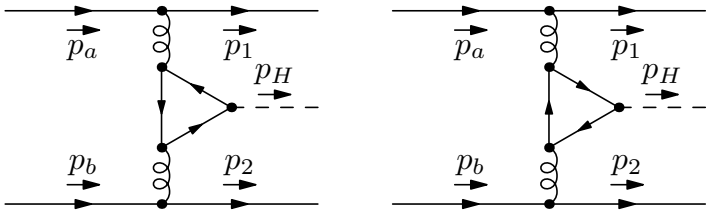
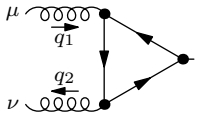



Figure 4.1: Leading-order diagrams contributing to the process $qQ \to qHQ$.

### 4.1.1 Finite mass dependence in $qQ \to qHQ$

We first consider the gluon-fusion production of a Higgs boson with two jets originating from two quarks with different flavours $q$ and $Q$. At leading order, only two diagrams contribute to this production channel for each quark flavour propagating in the loop, see fig. 4.1. As illustrated, we choose the momentum assignment $q(p_a)Q(p_b) \to q(p_1)H(p_H)Q(p_2)$ with all momenta left-to-right.

The quark-loop insertions in the two diagrams are symmetric under charge conjugation. They can be written as a colour-diagonal vertex of the form

$V_H^{\mu\nu}(q_1,q_2) =$ 

$= \dfrac{\alpha_s m^2}{\pi v}\Big[g^{\mu\nu}T_1(q_1,q_2) - q_2^\mu q_1^\nu T_2(q_1,q_2)\Big],$ (4.1.1)

where $m$ is the quark mass, and $v$ is the Higgs vacuum expectation value. For convenience, we list the form factors $T_1$ and $T_2$ in appendix B.1. The expressions are given there for a single propagating quark flavour; in practice, we sum the contributions from a top- and a bottom-quark propagating in the loop at the amplitude level. We therefore have contributions from both flavours *and* the interference between them when we square the amplitude. The colour and helicity summed and averaged matrix element then takes the factorised form

$$\overline{\left|\mathcal{M}_{qQ\to qHQ}\right|}^2 = \frac{1}{4(N_C^2-1)}||\mathcal{S}^m_{qQ\to qHQ}||^2 \cdot \left(g_s^2 C_F \frac{1}{t_1}\right)\cdot\left(\frac{1}{t_1 t_2}\right)\cdot\left(g_s^2 C_F \frac{1}{t_2}\right), \quad (4.1.2)$$

where the invariants $t_1 = q_1^2$, $t_2 = q_2^2$ are defined in terms of the $t$-channel momenta $q_1 = p_a - p_1, q_2 = p_2 - p_b$, and $\mathcal{S}^m_{qQ\to qHQ}$ is the current contraction

$$\mathcal{S}^m_{qQ\to qHQ} = j_\mu(p_1,p_a)V_H^{\mu\nu}(q_1,q_2)j_\nu(p_2,p_b), \quad (4.1.3)$$

where $j_\mu(p_o,p_i)$ is the generic $qq$-current from eq. (2.3.8).

The amplitude has the same structure as in the limit of an infinite quark mass eq. (2.3.16) [30–32]. Indeed, the only difference is in the expression for $V_H^{\mu\nu}$, which

in the limit agrees with eq. (2.3.17)

$$V_H^{\mu\nu}(q_1, q_2) \xrightarrow{m\to\infty} \frac{\alpha_s}{3\pi v}\left(g^{\mu\nu} q_1 \cdot q_2 - q_2^\mu q_1^\nu\right). \tag{4.1.4}$$

The factorisation of eq. (4.1.2) into a current contraction and a product of $t$-channel propagators is exactly what enables us to perform HEJ resummation. We stress that up to this point we retain the exact expression for the amplitude without having to resort to approximations valid in the high-energy limit.

Here we have referred to initial quarks; the treatment of initial antiquarks is completely analogous. The only qualitatively different amplitude is $q\bar{q} \to qH\bar{q}$, which receives contributions from both $t$-channel gluon exchange as in fig. 4.1 and two annihilation diagrams with $s$-channel gluon exchange. Both sets of diagrams are individually gauge independent, but the annihilation diagrams are subdominant for a large invariant mass between the two jets, so that the leading contribution to the amplitude is the same as the pure-quark amplitude eq. (4.1.2). One may also consider the process $q\bar{q} \to gHg$, but this is not an FKL configuration and hence will not contribute a leading power in $\log|\hat{s}/\hat{t}|$ to the matrix element.

### 4.1.2 Finite mass dependence in $gq \to gHq$

We already saw in section 2.3.1 for pure jets that the matrix elements for gluon-initiated $2 \to 2$ processes can also be described exactly as a contraction of currents, where the current for the equivalent quark process, i.e. $j_\mu^\pm(p_o, p_i)$ in eq. (4.1.3), is multiplied by a colour acceleration factor eq. (2.3.12). The $t$-channel factorisation of an amplitude implies not only that each factor is independent of the momenta of the rest of the process, but that it is also independent of the particle-content of the rest of the process. This description of incoming gluons in inclusive dijet production is therefore also valid in $H + 2j$ production and we will use it in the following.

At leading order, the $gq \to gHq$ subprocess is significantly more involved than the process with two incoming quarks. Of the 20 diagrams contributing to the LO amplitude, 10 can be obtained from charge conjugation. The remaining diagrams are depicted in fig. 4.2.

The amplitude with full quark-mass dependence is known for general kinematics [86, 161, 162]; it does not have a $t$-channel factorised form. In the following subsections, we discuss the different hierarchies which can exist between invariants in the process and the expressions we will use to describe this process in the corresponding regions of phase space.

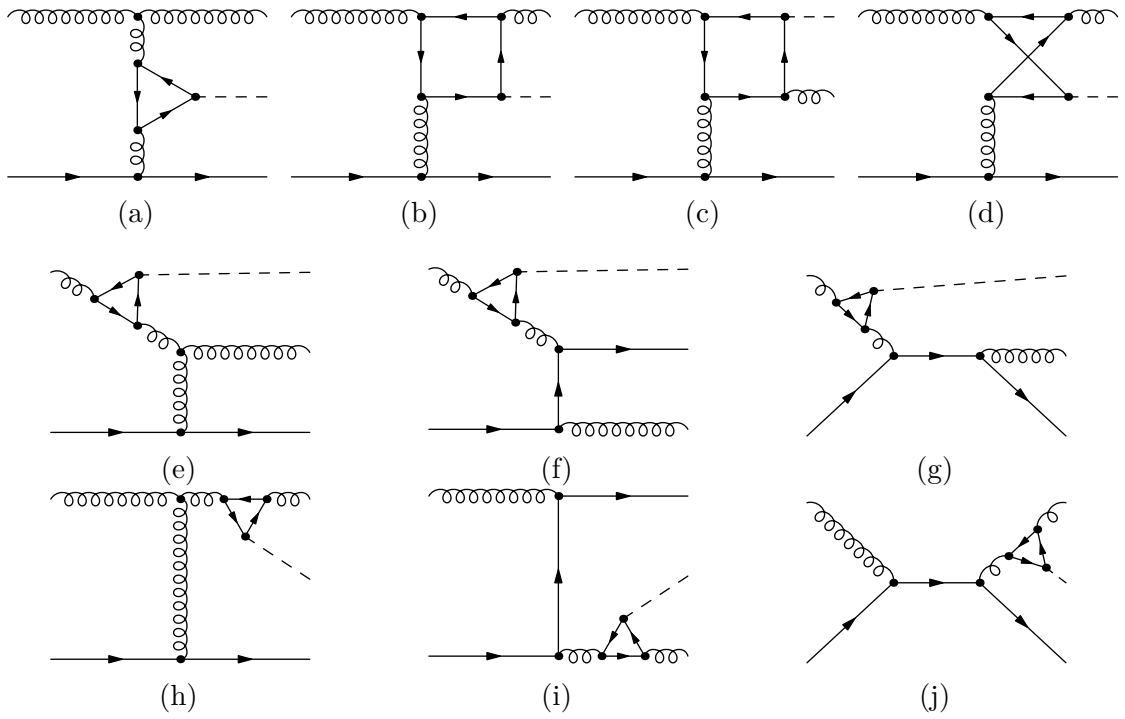



Figure 4.2: Leading-order diagrams contributing to the process $gq \to gHq$. Diagrams with clock-wise fermion flow in the heavy-quark loop can be obtained via charge conjugation and are not shown.

### Central Higgs-boson emission

Let us first discuss the case where the Higgs boson is emitted in between the two jets, with a large rapidity separation from each jet. More concretely, we consider the momentum assignment $g(p_a)q(p_b) \to g(p_1)H(p_H)q(p_2)$ with the hierarchy

$$s_{12} \gg s_{1H}, s_{2H} \gg t_1, t_2, m_H^2\,, \tag{4.1.5}$$

where $s_{ij} = (p_i + p_j)^2$ are invariant masses of the outgoing particles, $t_1 = (p_a - p_1)^2, t_2 = (p_b - p_2)^2$. Taking the gluon to be emitted backwards, this hierarchy implies the rapidity ordering $y_1 \ll y_H \ll y_2$. The forward emission of the gluon is of course completely analogous.

In ref. 109, it was shown that the amplitude in this limit assumes a similar factorised form as in the pure quark case. This is also true within the HEJ formalism [98]. As an example, the colour summed and averaged square of the helicity-conserving amplitude for a positive-helicity gluon and a negative-helicity quark can therefore be written as

$$\left|\mathcal{M}_{g^+q^-\to g^+Hq^-}\right|^2 = \frac{1}{N_C^2-1}||\mathcal{S}^m_{g^+q^-\to g^+Hq^-}||^2 \left(g_s^2 K_g(p_1^-, p_a^-)\frac{1}{t_1}\right)\cdot\left(\frac{1}{t_1 t_2}\right)\cdot\left(g_s^2 C_F \frac{1}{t_2}\right) \tag{4.1.6}$$

where $K_g(p_1^-, p_a^-)$ is the CAM factor from eq. (2.3.12). The current contraction

$$\mathcal{S}^m_{g^+q^-\to g^+Hq^-} = j^+_\mu(p_1, p_a)V_H^{\mu\nu}(q_1, q_2)j^-_\nu(p_2, p_b) \tag{4.1.7}$$

is otherwise completely equivalent to eq. (4.1.3) for $qQ \to qHQ$.

Similar to pure jets, the non-factorising helicity amplitudes involve a flip of the gluon helicity and identical helicities of the incoming partons [33]. These amplitudes are suppressed in the high-energy limit and therefore neglected here. Furthermore, independently of the chosen gauge, the contribution from the box diagrams, figs. 4.2b to 4.2d, is suppressed in the limit eq. (4.1.5). Thus for central Higgs-boson emission we can always use eq. (4.1.3)—up to the CAM factor for gluons.

### Peripheral Higgs-boson emission

We now consider the case where we drop the strong-ordering requirement between the Higgs boson and one of the jets. For the case where the Higgs boson is close or outside in rapidity of a quark, we again invoke $t$-channel factorisation to treat this as in the $qQ \to qHQ$ process. There the result with $V_H^{\mu\nu}$ is exact wherever the Higgs boson is emitted and hence in this case we again use eq. (4.1.6).

Thus we are left with only the case where the Higgs boson is not strongly ordered with respect to the rapidity of the gluon in $gq \to Hgq$. We will still require a separation from the quark, i.e. $y_1, y_H \ll y_2$ or

$$s_{12}, s_{2H} \gg s_{1H}, t_1, t_2, m_H^2\,. \tag{4.1.8}$$

We denote such configurations as $gq \to Hgq$, reserving the notation $gq \to gHq$ for the central Higgs-boson emission discussed in section 4.1.2. The $t$-channel factorisation of the amplitude is only guaranteed where there is a large rapidity separation between outgoing particles, and hence in this reduced limit we should not expect to recover a form with two $t$-channel poles as in eq. (4.1.6). However, as there is still a large rapidity separation to the quark line, we expect to find a factorised form about the pole in $t_2$ as follows

$$\overline{\left|\mathcal{M}_{gq\to Hgq}\right|}^2 = \frac{1}{4(N_C^2-1)}||\mathcal{S}^m_{gq\to Hgq}||^2 \cdot (\alpha_s^2 g_s^2 C_A) \cdot \left(\frac{1}{t_2}\right) \cdot \left(g_s^2 C_F \frac{1}{t_2}\right), \tag{4.1.9}$$

where the remainder of the amplitude

$$\mathcal{S}^m_{gq\to Hgq} = j_H^\mu(p_1, p_H, p_a)j_\mu(p_2, p_b) \tag{4.1.10}$$

has been written as an effective current $j_H^\mu$. This current, dependent on the reduced set of momenta $(p1, p_H, p_a)$, is derived in appendix B.2. The derivation follows closely

the approach in ref. 33 for calculating the effective gluon current. In contrast to that case, or indeed that of a central Higgs boson, amplitudes flipping the gluon helicity also contribute.

### 4.1.3 Finite mass dependence in $gg \to gHg$

As noted above, the $t$-channel factorisation which arises in the limit of large rapidity separations implies that the building blocks corresponding to each end of the chain is independent of the rest of the process. We can therefore describe the $gg$-initiated state by taking the expressions for $gq \to gHq$ and adding the necessary change for an incoming gluon derived from pure jets. We find for central Higgs-boson emission from eq. (4.1.6):

$$\begin{aligned}\left|\mathcal{M}_{g^+g^-\to g^+Hg^-}\right|^2 =& \frac{1}{N_C^2-1}||\mathcal{S}^m_{g^+q^-\to g^+Hq^-}||^2 \\ &\cdot\left(g_s^2 K_g(p_1^-,p_a^-)\frac{1}{t_1}\right)\cdot\left(\frac{1}{t_1t_2}\right)\cdot\left(g_s^2 K_g(p_2^+,p_b^+)\frac{1}{t_2}\right)\end{aligned} \tag{4.1.11}$$

$$\mathcal{S}^m_{g^+g^-\to g^+Hg^-} = j^+_\mu(p_1,p_a)V_H^{\mu\nu}(q_1,q_2)j^-_\nu(p_2,p_b). \tag{4.1.12}$$

Likewise, for a Higgs boson emitted backward of both gluons where only the subset of hierarchies applies (eq. (4.1.8)), we find from eq. (4.1.9):

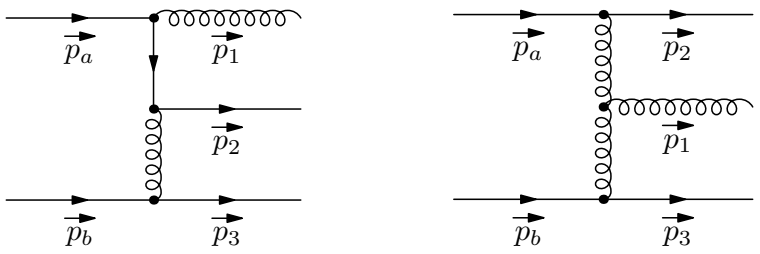
$$\overline{\left|\mathcal{M}_{gg\to Hgg}\right|}^2 = \frac{1}{4(N_C^2-1)}||\mathcal{S}^m_{gg\to Hgg}||^2\cdot(\alpha_s^2 g_s^2 C_A)\cdot\left(\frac{1}{t_2}\right)\cdot\left(g_s^2 K_g(p_2^+,p_b^+)\frac{1}{t_2}\right), \tag{4.1.13}$$

$$\mathcal{S}^m_{gq\to Hgq} = j_H^\mu(p_1,p_H,p_a)j_\mu(p_2,p_b). \tag{4.1.14}$$

Note that the gluon closest to the Higgs boson will always have the more complicated treatment derived in the previous subsection, i.e. it is the momentum of the gluon closest in rapidity to the Higgs boson which will enter $j_H^\mu$.

### 4.1.4 The first set of next-to-leading logarithmic corrections

We have now assembled the HEJ description of the scattering amplitudes for all Higgs boson plus (exactly) two jets processes which will contribute at leading power for Higgs-plus-dijets. At the start of this section, we identified the necessary subprocesses as the FKL configurations of flavour and momenta. In ref. 98, the HEJ framework was extended to also describe Born processes where the requirement of rapidity ordering on exactly one gluon was relaxed, allowing it to be emitted outside of the




$y_1 < y_2 < y_3$ $\qquad$ $y_2 < y_1 < y_3$

Figure 4.3: Rapidity-ordered ladder diagrams for $qQ \to gqQ$. The ordering in eq. (4.1.15) (left) contains just one $t$-channel gluon propagator, whereas the FKL configuration (right) contains the maximum number, two.

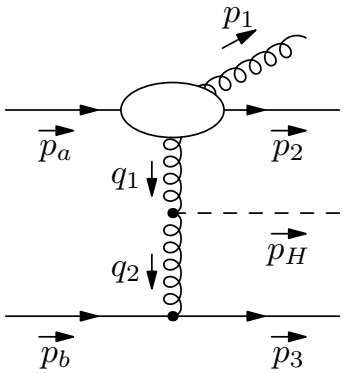



Figure 4.4: For the rapidity ordering in eq. (4.1.15) where there is only strong ordering between $y_2$ and $y_3$, we only find factorisation about the $t$-channel pole between these particles so the structure of the amplitude are as shown in eq. (4.1.17).

rapidity range defined by an outgoing quark, e.g.

$$q(p_a)Q(p_b) \to g(p_1)q(p_2)H(p_H)Q(p_3) \qquad y_1 < y_2 \ll y_3\,. \tag{4.1.15}$$

We already saw in section 2.2 that this corresponds to an unordered configuration, which is suppressed by one power of $\log|\hat{s}/\hat{t}|$. This is formally therefore a next-to-leading logarithmic contribution to the dijet cross section; however one can still construct the leading logarithmic contributions to each particular subprocess. This particular class of processes was chosen as it had been observed after matching to leading order that they contributed significantly in regions of phase space with large transverse momentum. Their inclusion therefore allows HEJ to reduce its dependence on fixed-order matching [98].

From eq. (4.1.15), these subprocesses at Born-level have just one strong rapidity-ordering between the coloured particles and one therefore constructs an effective

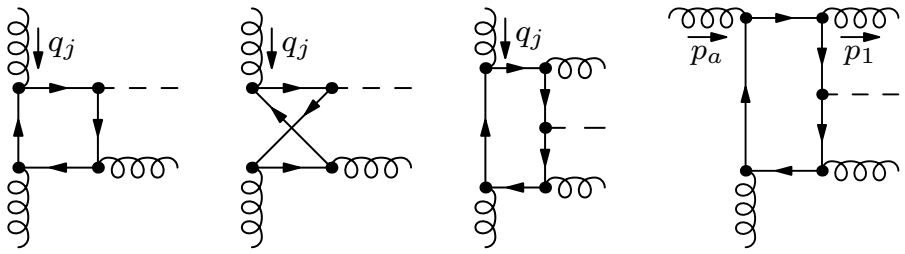



Figure 4.5: Examples for additional gluon emission off the heavy-quark loop. These emissions are suppressed for large rapidity separations. Here vertical lines represent off-shell $t$-channel propagators and horizontal lines represent on-shell external particles.

current to describe the $q(p_a) \to g(p_1)q(p_2)g^*$ end of the chain, denoted $j_\mu^{\text{uno}}(p_2, p_1, p_a)$ (see fig. 4.4). This must now carry two colour indices as it consists of terms with different colour flow. The matrix element for the case of central Higgs-boson emission—in which the Higgs boson is emitted between *the quarks*—is then given by

$$\begin{aligned}\overline{\left|\mathcal{M}_{qf_2\to gqHf_2}\right|}^2 = &\frac{1}{4(N_C^2-1)}||\mathcal{S}^{\text{uno}}_{qf_2\to gqHf_2}(p_1,p_2,p_3,p_a,p_b,q_1,q_2)||^2 \\ &\cdot\left(g_s^4 K_{\text{uno}}\frac{1}{t_1}\right)\cdot\left(\frac{1}{t_1 t_2}\right)\cdot\left(g_s^2 K_{f_2}\frac{1}{t_2}\right),\end{aligned} \tag{4.1.16}$$

$$\mathcal{S}^{\text{uno}}_{qf_2\to gqHf_2}(p_1,p_2,p_3,p_a,p_b,q_1,q_2) = \frac{1}{\sqrt{C_F}} j_\mu^{\text{uno }cd}(p_2,p_1,p_a)V_H^{\mu\nu}(q_1,q_2)j_\nu(p_3,p_b)T^c_{3b}\,, \tag{4.1.17}$$

where $K_{\text{uno}} = -1/2$ and we are using $q_1 = p_a - p_1 - p_2$, $q_2 = p_3 - p_b$, $t_i = q_i^2$ (as in fig. 4.4). It is clear that the arguments of $\mathcal{S}^{\text{uno}}_{qQ\to gqHQ}$ are not independent; we have chosen to display the implicit dependence on $q_1, q_2$ in anticipation of processes with additional gluons, where the dependence is explicit. The expression for $j_\mu^{\text{uno}}(p_2, p_1, p_a)$ is given in appendix B.3.

With this we have constructed the HEJ approximations to Born-level matrix elements. For the generalisation to arbitrary multiplicity we can follow the procedure outlined in section 2.3. The only (potential) complication comes from the emission of an extra gluon from the heavy quark loop in the Higgs-boson coupling (see fig. 4.5). Similar to the discussion for central Higgs-boson emission, the box diagrams are suppressed at large rapidity separation between the gluon and the Higgs boson. The absence of $t$-channel enhancement is obvious in the limit of a large quark mass, where the quark loop is absorbed into an effective local interaction. Thus, we can neglect these terms; the remaining effective Lipatov vertex is then exactly that of eq. (2.3.25). Consequently, all discussion about the virtual corrections from section 2.3.3 apply.

The final regularised, resummed expressions for the matrix element for the production

of a Higgs boson in between partons $j$ and $j+1$ with incoming flavours $f_1, f_2$ is then[3]

$$\begin{aligned}\overline{\left|\mathcal{M}^{\mathrm{HEJ}}_{f_1 f_2 \to f_1 \cdot g \cdot H \cdot g \cdot f_2}\right|}^2 = \frac{1}{4(N_C^2-1)} \left\|\mathcal{S}_{f_1 f_2 \to f_1 H f_2}\right\|^2 \\ \cdot \left(g_s^2 K_{f_1}(p_1^-, p_a^-)\frac{1}{t_1}\right) \cdot \left(\frac{1}{t_j t_{j+1}}\right) \cdot \left(g_s^2 K_{f_2}(p_n^+, p_b^+)\frac{1}{t_n}\right) \\ \cdot \prod_{k=2}^{j} \left(\frac{-g_s^2 C_A}{t_{k-1}t_k} V_L^{\nu_k}(q_{k-1}, q_k) V_{L\nu_k}(q_{k-1}, q_k)\right) \\ \cdot \prod_{k=j+1}^{n-1} \left(\frac{-g_s^2 C_A}{t_k t_{k+1}} V_L^{\nu_k}(q_k, q_{k+1}) V_{L\nu_k}(q_k, q_{k+1})\right) \\ \cdot \prod_{i=1}^{j-1} \exp\left[\omega^0(q_{i\perp})(y_{i+1}-y_i)\right] \cdot \prod_{i=j+2}^{n} \exp\left[\omega^0(q_{i\perp})(y_i - y_{i-1})\right] \\ \cdot \exp\left[\omega^0(q_{j\perp})(y_H - y_j)\right] \cdot \exp\left[\omega^0(q_{j+1\perp})(y_{j+1} - y_H)\right],\end{aligned} \tag{4.1.18}$$

where $\cdot g \cdot$ stands for arbitrarily many gluon emissions.

The same formula holds for a backward (forward) Higgs-boson emission if $f_1$ ($f_2$) is a quark or antiquark with $j = 1$ ($j = n$). For a peripheral emission close to a gluon, there is an equivalent expression but the first two lines instead mirror eq. (4.1.9) and we have one fewer $t$-channel pole as there is one fewer hierarchy. Explicitly, the matrix-element-squared is given by

$$\begin{aligned}\overline{\left|\mathcal{M}^{\mathrm{HEJ}}_{g f_2 \to H g \cdot g \cdot f_2}\right|}^2 = \frac{1}{4(N_C^2-1)} \left\|\mathcal{S}^m_{g f_2 \to H g f_2}\right\|^2 \cdot (\alpha_s^2 g_s^2 C_A) \cdot \left(\frac{1}{t_1}\right) \cdot \left(g_s^2 K_{f_2}(p_n^+, p_b^+)\frac{1}{t_n}\right) \\ \cdot \prod_{k=2}^{n-1} \left(\frac{-g_s^2 C_A}{t_{k-1}t_k} V_L^{\nu_k}(q_{k-1}, q_k) V_{L\nu_k}(q_{k-1}, q_k)\right) \\ \cdot \prod_{i=1}^{n-1} \exp\left[\omega^0(q_{i\perp})(y_{i+1}-y_i)\right].\end{aligned} \tag{4.1.19}$$

Note here that $q_1 = p_a - p_1 - p_H$, $q_i = q_{i-1} - p_i$ for $i = 2, \ldots, n$ and $t_i = q_i^2$.

The regulated matrix element above is valid in the phase space of an arbitrary number of extra real gluon emissions each with $|p_\perp| > \kappa$, provided they are between the extremal partons in rapidity. Note that the extremal partons play a special role and are not allowed to become soft, since we do not include the necessary virtual corrections to regulate the fundamental spinor strings. In practice we require the extremal partons to carry a significant fraction of the extremal jet momentum to ensure that they remain perturbative.

[3]Note that the factor of $1/(t_j t_{j+1})$ in the second line was missing in ref. 98.

## 4.2 Matching to fixed order

This section describes the matching of results within HEJ both to the full leading-order finite quark-mass matrix elements and to the NLO cross sections obtained for infinite top-mass. We generate our event weights using the procedure outlined in chapter 3, where we begin from fixed-order samples and supplement these with resummation. The resummation step only applies to the particle and momentum configurations discussed in section 4.1, namely FKL configurations or pseudo-FKL with one unordered gluon. If a given fixed-order event is not one of these configurations, it enters our final event sample with its weight unaltered.

We observed in section 3.2 that a central scale choice of $\mu_r = \mu_f = H_T/2$ leads to distributions in rapidity and dijet invariant mass with values close to the upper edge of the scale variation band obtained when $\mu_r$ and $\mu_f$ are varied independently by a factor of two around this central scale choice, keeping their ratio between 0.5 and 2. The scale variance obtained with a central scale of $\mu_r = \mu_f = H_T/2$ is therefore pathologically small for distributions at large dijet invariant mass or large rapidity separations, which are relevant for the VBF studies. While the scale variations obtained at NLO with a central scale choice of $\mu_r = \mu_f = \max(m_H, m_{12})$ are larger, they are also more reasonable as an indication of the uncertainty due to higher order corrections within the VBF-cuts. We therefore use $\mu_r = \mu_f = \max(m_H, m_{12})$ as central scale in this chapter. We note that, with this choice, the $p_\perp$-based observables such as $p_{H\perp}$ show the same pathological scale variance for large values that the invariant mass-based observable develops for renormalisation and factorisation scales based on the transverse momenta.

Section 4.2.1 describes the fixed-order samples available which we use as our starting point and the point-by-point matching applied to the resummation events. Section 4.2.2 then describes the matching performed for all events at the level of the total cross section.

### 4.2.1 Matching of exclusive amplitudes

HEJ allows the perturbative series for each $n$-jet phase space point to be matched to fixed order. This obviously is possible only if amplitudes for the $n$-jet phase space point are readily available. In this study, the fixed-order calculations are performed using Sherpa interfaced with OpenLoops for the evaluation of the $pp \to H + 2j$-processes with full quark-mass dependence. This fixed-order setup includes just the effects from the top-quark, and not also those of the loops of bottom-quarks. The

effect of both top- and bottom-mass is included in the resummation, and will be discussed later.

Additionally, the $H + 3j$-processes are unavailable with the full quark-mass dependence even at leading order, and even using the infinite top-mass limit, only the $pp \to H + 2j$-process is available at NLO (and therefore $pp \to H + 3j$ at tree-level).

The limitations in the fixed-order results mean that the matching within HEJ has to use a large number of components than in chapter 3. We will describe them in the following sections. No point-by-point matching is performed for events with six or more jets.[4] For such multiplicities the fixed-order results are expensive to compute, while typically only contributing less than a percent for all shown observables. Instead, similar to the previous chapter, we use HEJ FOG to approximate these LO results.

On top of the matching of exclusive events described in the following, the final predictions for HEJ will be scaled with the ratio of the inclusive cross section for $pp \to H + 2j$ calculated at infinite top-mass for NLO and HEJ.

The described procedure obtains top- and bottom-mass dependence through the all-order results, matching to the full top-mass results for $pp \to H + 2j$, and to $pp \to H + 3j$, $H + 4j$ and $H + 5j$ in the limit of infinite top-mass.

### Two-jet matching with finite quark mass

The exclusive two-jet events are matched to full leading order, with finite quark mass effects. However, as our fixed-order setup allows for just the top-quark diagrams, technically the matching is performed by multiplying the all-order results containing both top- and bottom-mass effects with the ratio of the square of the full Born-level matrix element evaluated with just the top-quark and the corresponding approximation within HEJ (using just the propagating top-quark, with no bottom-quark effects). The final event weights are therefore proportional to

$$|\mathcal{M}_{\text{HEJ}}^{m_t,m_b}|^2 \frac{|\mathcal{M}_{\text{LO}}^{m_t}|^2}{|\mathcal{M}_{\text{HEJ,LO}}^{m_t}|^2}\,, \tag{4.2.1}$$

where $|\mathcal{M}_{\text{LO}}|^2$ is the leading-order matrix element, $|\mathcal{M}_{\text{HEJ}}|^2$ the all-order HEJ matrix element, and $|\mathcal{M}_{\text{HEJ,LO}}|^2$ its truncation to leading order. The superscript indicates the quark masses that are taken into account.

[4] We therefore match one more jet compared to section 3.2.3.

### Three-, four- and five-jet matching with infinite quark mass

With the method of chapter 3, the resummation could be constructed starting from event files from the calculation of Born-level Higgs-boson production in association with three jets including full momentum and mass dependence reported in ref. 164. However, since these are not available, the three-, four- and five-jet events will be matched in the infinite top-mass limit, which can be readily evaluated using Sherpa and Comix [138]. Technically then, the reweighting of the event is performed with the ratio of the Born-level evaluation of the HEJ-approximation in the infinite top-mass and the full Born-level expression in the same limit, while the resummation is performed using the full expressions developed in section 4.1, including top- and bottom-mass. The contribution from the matrix elements to the event weights is then

$$|\mathcal{M}_{\text{HEJ}}^{m_t,m_b}|^2 \frac{|\mathcal{M}_{\text{LO}}^{\text{eff}}|^2}{|\mathcal{M}_{\text{HEJ,LO}}^{\text{eff}}|^2}\,, \tag{4.2.2}$$

where the "eff" superscript refers to the effective theory of an infinite top-quark mass. In this approximation, the interaction between the Higgs boson and gluons is described by an operator of dimension five, so that matrix elements exhibit unphysical scaling in the limit of large momenta. Since we choose not to include finite top-mass corrections in the truncated HEJ matrix element this effect cancels out in the ratio in eq. (4.2.2).

The emission of quarks and gluons should resolve the dependence on the loop momenta only for large energies of the emission with respect to $m_t$. Since the bulk of each jet multiplicity consists of jet transverse momenta close to the defined jet threshold, the quark-mass effects should have only a small effect on the inclusive cross section. The quality of the approximation can be checked by applying a similar strategy of reweighting in $pp \to H + 2j$, where the full expression is known and can be used as a benchmark.

In fig. 4.6 we show the results for the different approximations of $pp \to H + 2j$ at Born-level. When assuming $m_t \to \infty$ the simulation undershoots the full finite top-mass calculation by 5% for transverse momenta of the Higgs boson up to $p_{H\perp} \approx m_t$. At even higher values it diverges from the cross section with full top mass dependence. While the corrections are relatively small and uniform for the differential cross section with respect to the azimuthal angle between the two jets and the rapidity of the Higgs boson, there are systematically increasing corrections to the distribution with respect to the invariant mass between the two jets, growing to more than 10% for $m_{12} > 700\,\text{GeV}$. The underestimation of the peak, and the overestimation in the tail cancel, such that the inclusive cross section shows good agreement between finite and infinite top-mass, as observed in ref. 109, 164. However, this cancellation is

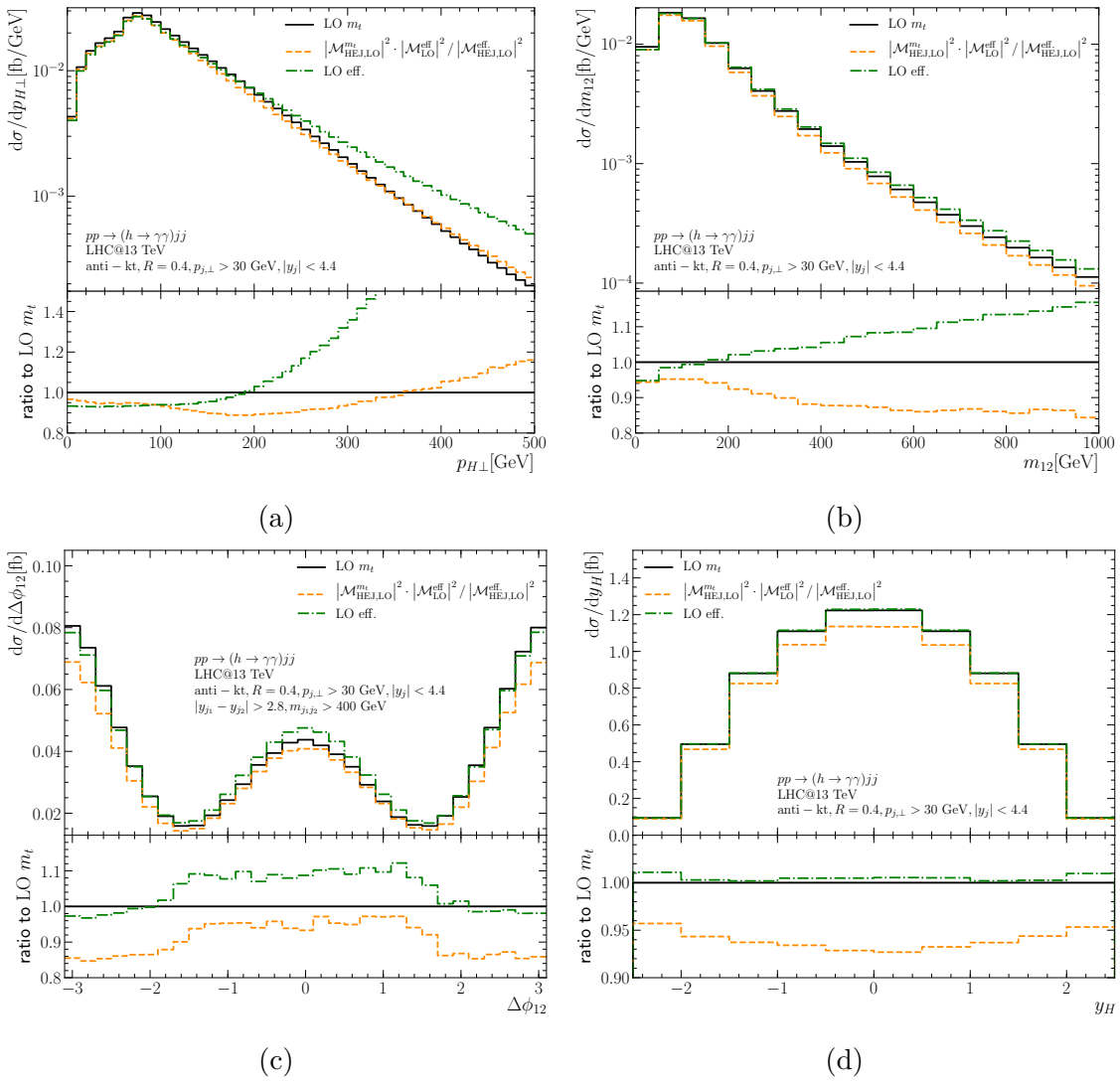



Figure 4.6: In these plots we check the quality of reweighting using a ratio of matrix elements with infinite quark mass, using $pp \to H + 2j$. The black, solid line shows the LO calculation with full top-mass dependence and the green, dot-dashed line shows LO in the effective theory. The orange, dashed line shows the HEJ result truncated at Born-level with full top-mass dependence, reweighted with the ratio of LO to HEJ matrix elements in the infinite top-mass limit (see eq. (4.2.2)). The deviation between the orange and black lines arises from the infinite top-mass limit in the reweighting factor.

accidental and will depend on the transverse cuts used.

We further compare the pure LO calculations to an approximation where a finite top-mass comes only from the HEJ Born-level matrix element. For this, we multiply the infinite top mass LO matrix element with the ratio of Born-level HEJ at finite over infinite top mass. This is similar to the procedure eq. (4.2.2) which we use for 3, 4 and 5 jets. If HEJ would catch all finite top-mass effects for all processes, the solid (black) and dashed (orange) lines in fig. 4.6 would be identical; on the other hand if HEJ would not describe any top-mass effect the dashed (orange) and the dot-dashed (green) line would agree. Thus the difference gives a measure of the quality of the approximation. We see that for rapidity-distributions, exemplified by that of the Higgs boson, the level of accuracy obtained is roughly 5%. The accuracy is better than 12% in the distribution of the azimuthal angle between the two jets, and similar for the invariant mass between the jets, although here, and for the transverse mass-distribution of the Higgs boson, the corrections increase with increasing scale.

We emphasise that the seemingly good agreement in the distribution of $y_H$ between the results using the infinite top-mass and full top-mass dependence is completely accidental. The normalisation only agrees since $y_H$ effectively integrates over $p_{H\perp}$. The results presented with the dashed (orange) line and obtained using eq. (4.2.2) are more accurate.

We conclude that by using finite quark-masses in the simplified HEJ amplitudes, and applying matching in the infinite top-mass limits we can expect the result with finite top- (and bottom-)quark mass to be within 20% for all distributions. This is sufficiently accurate for 3, 4 and 5 jets.

### 4.2.2 Matching of LO to NLO in the infinite quark mass limit

The resummation and matching procedure described so far will be compared to the best possible fixed-order result we can obtain. This consists of Born-level for full top-mass, but not including the small effect of the bottom-mass, reweighted bin-by-bin by the differential NLO $K$-factor calculated for infinite top mass. The LO and NLO calculations for the distribution of (left) the rapidity separation of the hardest two jets and (right) the maximum rapidity-difference between any two hard jets, $\Delta y_{fb}$ in $pp \to H + 2j$ with infinite top-mass are shown in fig. 4.7.

The NLO $K$-factor is particularly interesting: it has a linear growth in both cases and is large at large $\Delta y$. Where it is plotted as function of $\Delta y_{fb}$, it goes to 1 for

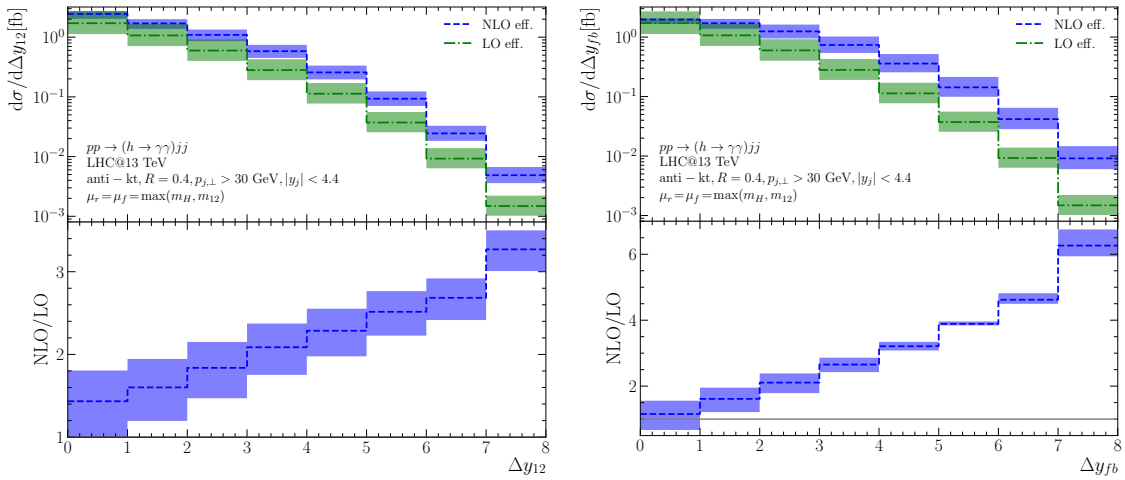



Figure 4.7: The distribution of the rapidity separation between the hardest jets (left) and the most forward and backward jets (right) of Higgs-plus-dijet production for LO (green, dot-dashed) and NLO (blue, dashed) both in the infinite top-mass limit. The bottom panel shows the NLO $K$-factor in each case. The calculations where done with a central scale $\mu_r = \mu_f = \max(m_H, m_{12})$.

$\Delta y_{fb} = 0$. This represents a region of phase space dominated by exclusive 2-jet events. For the rapidity separation of the hardest two jets the $K$-factor reaches a factor of 3 for rapidity differences of $\Delta y_{12} = 8$, and for the most forward/backward jets reaches a factor of 6 at $\Delta y_{fb} = 8$. This obviously brings into question the validity of NLO-calculations at such rapidity-differences. This source of the apparent perturbative instability in the fixed-order result is treated systematically within HEJ.

It is worth mentioning here that in the MRK-limit the all-order HEJ-resummation for $pp \to H + 2j$ and $pp \to H + 3j$ will contain the same effect from the virtual corrections (and soft emissions) of a suppressing factor $\exp(\omega(k_\perp^2)\Delta y_{fb}) \propto (\hat{s}/p_\perp^2)^{\omega(k_\perp^2)}$ with $\omega(k_\perp^2) < 0$. However, when the perturbative series is terminated at NLO-accuracy, the effect of the expansion of the exponential suppression is included only in the events with Born-level kinematics. The suppression is missing at NLO in the corrections from real emissions because of the fixed-order termination of the perturbative series. At large rapidity-spans $\Delta y_{fb}$, this will inflate the NLO predictions compared to the all-order result of HEJ, irrespectively of the choice of renormalisation and factorisation scale.

The balance between a suppression for $H + 2j$ at NLO at large $\Delta y$ of the two-parton contribution and enhancement of the three-parton contribution discussed above is obviously influenced by the value of $\alpha_s$ and therefore the scale choices. Indeed, the effect of choosing instead a central scale choice of $\mu_r = \mu_f = H_T/2$ is illustrated in fig. 4.8. As seen in the right plot, the $K$-factor tends to unity for $\Delta y_{fb} \to 0$,

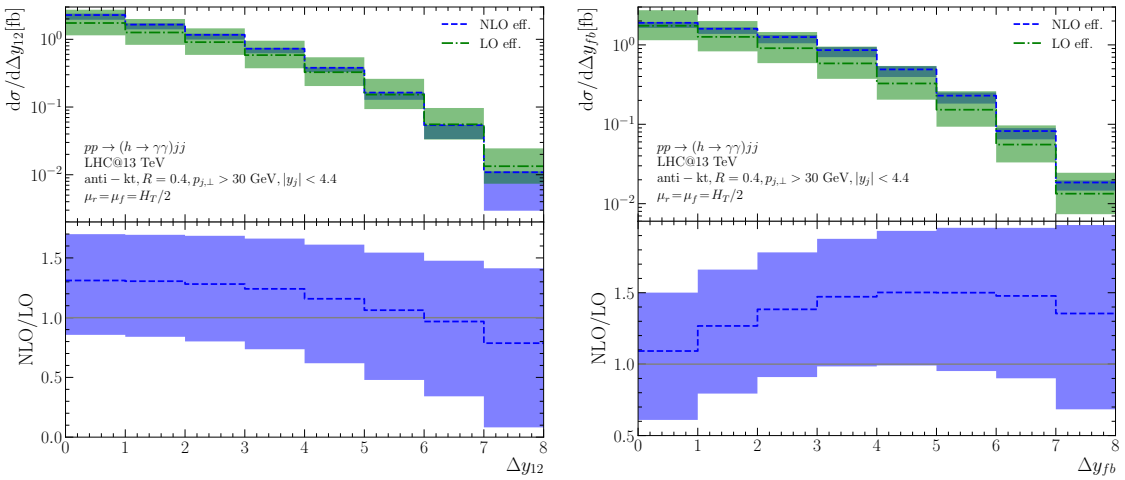



Figure 4.8: The distribution of the rapidity separation between the hardest jets (left) and the most forward and backward jets (right) of Higgs-plus-dijet production. Setup is the same as fig. 4.7 but with $\mu_r = \mu_f = H_T/2$.

rises to 1.5 at $\Delta y_{fb} = 4$, stabilises and then starts decreasing at $\Delta y_{fb} \sim 7$. As a function of $\Delta y_{12}$, the $K$-factor starts at 1.3 for $\Delta y_{12} = 0$, and then decreases to 0.7 at $\Delta y_{12} = 8$. The smaller $K$-factors observed for the central scale choice of $\mu_r = \mu_f = H_T/2$ may seem more appealing than the behaviour observed in fig. 4.7; however, the variation obtained around this central scale will certainly underestimate the uncertainty from uncalculated higher orders, since the central scale choice leads to a result close to the edge of the results obtained by the variation. Furthermore, the scale variation band for NLO in fig. 4.8 (left) increases with $\Delta y_{12}$, reaching $-70\%$ in the last bin, above $\Delta y_{12} > 7$. This indicates an instability of the NLO calculations for $\mu_r = \mu_f = H_T/2$ at large rapidity differences. All the results presented in the following with $\mu_r = \mu_f = \max(m_H, m_{12})$ are also presented in appendix C for $\mu_r = \mu_f = H_T/2$. Just as for fixed-order predictions, other processes like $W$ with jets could be used in order to verify which of the scale choices obtains the best description of data.

## 4.3 Results for finite quark masses and all-order resummation

This section will first present results for a separate investigation of the higher-order effects included with HEJ compared to the fixed-order approaches. As mentioned before, we employ Sherpa in combination with OpenLoops to obtain the fixed-order predictions. To evaluate the finite quark-mass corrections within HEJ, we

make use of QCDLoop [165]. The input parameters and cuts are (almost) identical to the ones described in section 3.2.1. In particular, we use the cuts from eq. (3.2.1) for the inclusive cross section, and eq. (3.2.2) as the VBF selection. Unlike the previous chapter, we choose a central scale of $\mu_r = \mu_f = \max(m_H, m_{12})$, without special treatment of the Higgs coupling; all figures in this section are reproduced in appendix C for the central scale of $\mu_r = \mu_f = H_T/2$.

A discussion of the values chosen for the quark masses is in order. In the gluon-fusion production of a Higgs boson together with light-flavour jets the heavy quarks only appear in internal loops and are off-shell. We therefore do not use on-shell masses, but instead prefer the $\overline{\text{MS}}$ mass-scheme. The scale $\mu_m$ associated with the $\overline{\text{MS}}$ mass is a priori independent of the renormalisation scale used for the running coupling. It should be set to a scale characteristic for the heavy-quark loop.

For the bottom-quark, the mass is negligible compared to all other scales in the loop. Since the observables considered in this work depend only mildly on the bottom-quark mass, the exact scale choice has little impact on the prediction. To be definite, we use $\mu_{m_b} = m_H$ and $m_b(m_H) = 2.8\,\text{GeV}$, which can be obtained from input values of $m_b(m_b) = 4.18\,\text{GeV}$ [57], $\alpha_s(m_Z) = 0.118$ via renormalisation group evolution at two loops. The effect of higher orders in the evolution is negligible.

The effect of the top-quark mass is much more important. While there are ongoing efforts [166–171] to relate the very precise values reported by the LHC and Tevatron experiments [172–174] to a well-defined short-distance scheme, the top-quark $\overline{\text{MS}}$ mass is not known very precisely at the moment. For this project, the values chosen are $\mu_{m_t} = m_H$ with $m_t(m_H) = 163\,\text{GeV}$, in line with direct determinations of the $\overline{\text{MS}}$ mass [175, 176] and compatible with a pole mass of 173 GeV [177] within the uncertainties quoted in ref. 176.

Since the fixed-order setup can take into account the effects only of the top-quark, all simulations in section 4.3.1 are for finite top-mass only (no effects from the bottom-quark included). Section 4.3.2 investigates the effects on the HEJ resummation with finite top and bottom mass compared to infinite top mass. Finally, section 4.3.3 compares the most precise predictions from HEJ, including both top- and bottom-mass effects, and the matching to fixed order discussed in section 4.2, to that of fixed order with finite top-mass, scaled to NLO accuracy, as described in section 4.2.2.

### 4.3.1 Effects of higher perturbative orders

Figure 4.9 compares the results obtained with finite top-mass at LO, the LO predictions rescaled to NLO accuracy in the limit of infinite top-mass, and with all-order

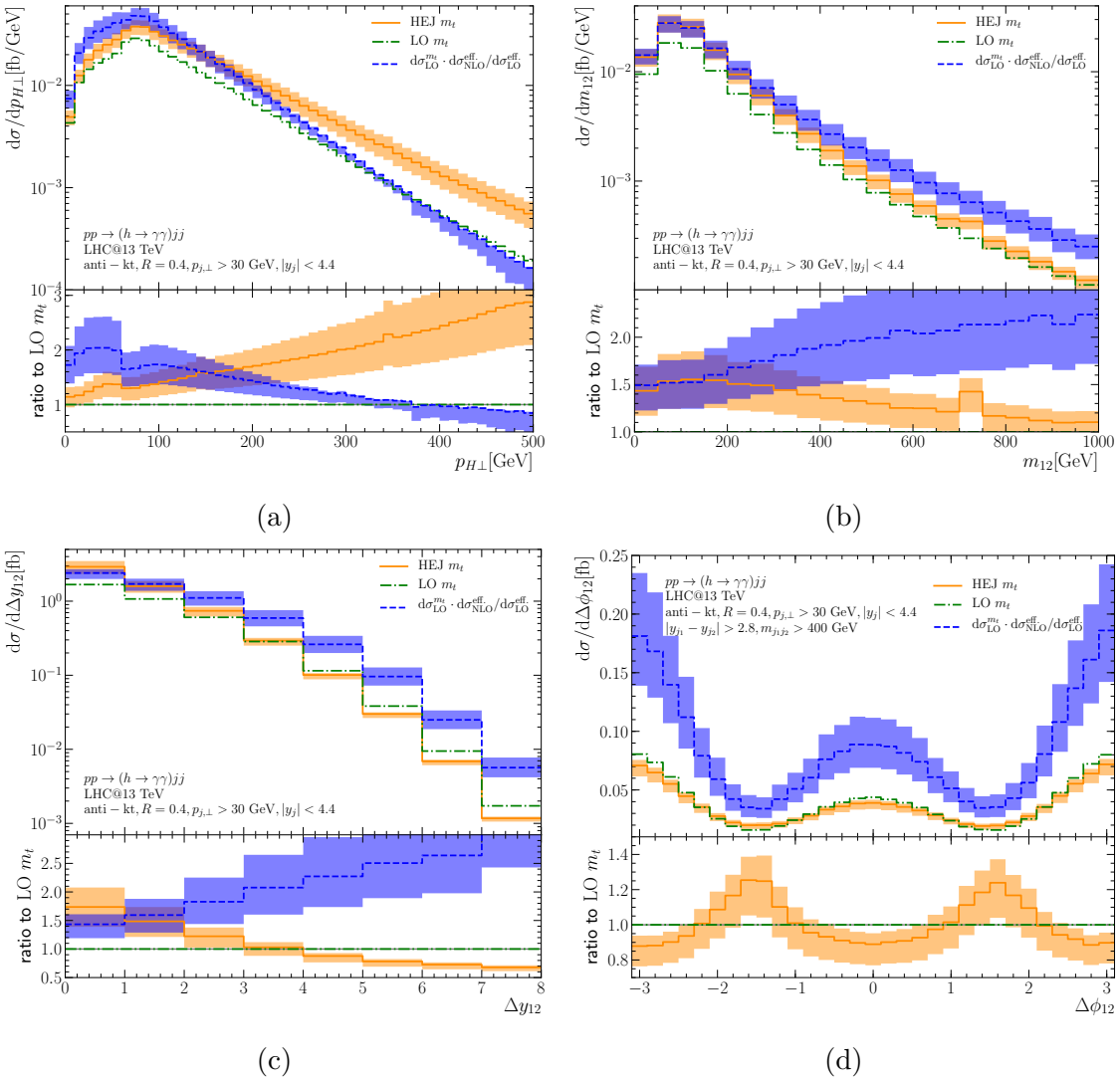



Figure 4.9: Differential cross section for $p_{H\perp}$ (a), $m_{12}$ (b), $\Delta y_{12}$ (c) and $\Delta\phi_{12}$ (d) obtained from HEJ (orange, solid) and LO (green, dot-dashed) both with full top-mass. Additionally in blue, dashed is the result of scaling LO full $m_t$ bin-by-bin with the NLO $K$-factor in the $m_t \to \infty$ limit. The $K$-factors and their impact within the VBF cuts (applied in (d)) are discussed in the text.

| Distribution | LO($m_t$) | LO($m_t$) $*$ $K_{\mathrm{NLO}}(m_t \to \infty)$ | HEJ($m_t$) |
|---|---|---|---|
| Inclusive | $3.8^{+2.1}_{-1.3}$ fb | $6.2^{+1.1}_{-1.2}$ fb | $5.7^{+1.0}_{-1.1}$ fb |
| $p_{H\perp}$ | | $6.3^{+1.2}_{-1.3}$ fb | |
| $m_{12}$ | | $6.2^{+1.0}_{-1.1}$ fb | |
| $y_{12}$ | | $6.2^{+1.1}_{-1.2}$ fb | |
| VBF | $0.24^{+0.12}_{-0.08}$ fb | $0.53^{+0.15}_{-0.13}$ fb | $0.23^{+0.02}_{-0.03}$ fb |
| VBF, $\phi_{12}$ | | $0.53^{+0.13}_{-0.10}$ fb | |

Table 4.1: Cross sections obtained at LO, LO scaled with bin-by-bin $K$-factor for various distributions, and the HEJ with the inclusive cross sections scaled to NLO

HEJ (using just finite top-mass but no contribution from the bottom-quark). Comparing the cross section over $p_{H\perp}$ in fig. 4.9a for LO and the rescaling, using the bin-by-bin $K$-factor calculated in the limit of infinite top-mass, one sees that the NLO $K$-factor (the ratio between the lines in blue and in green, indicated by the blue band in the lower plots) varies locally between 0.8 and 2 within ranges of the distributions checked. The NLO $K$-factor is decreasing for increasing transverse momentum $p_{H\perp}$, crossing unity at $p_{H\perp} = 340\,\mathrm{GeV}$.

The NLO $K$-factors for the distributions in the invariant mass between the two hardest jets $m_{12}$ (fig. 4.9b) and the rapidity-difference between the two hardest jets (fig. 4.9c) have the same systematic behaviour of increasing $K$-factor as observed for $\Delta y_{fb}$ in fig. 4.7 and discussed there. The NLO $K$-factor for $m_{12}$ increases from 1.5 to 2.2 at $m_{12} = 1\,\mathrm{TeV}$, and for $\Delta y_{12}$ the NLO $K$-factor increases in a straight line from 1.5 to 3 at $\Delta y_{12} = 8$. This obviously then induces a large $K$-factor when a large rapidity-separation and invariant mass is required in the VBF-cuts, as illustrated for $\phi_{12}$ within these plots seen in fig. 4.9d. It can also be seen in fig. 4.9c that the ratio between HEJ and LO decreases linearly as a function of $\Delta y_{12}$; this is an illustration of the logarithmic suppression of events with exactly two jets where $\Delta y = \log|\hat{s}/\hat{t}|$ for large $\hat{s}$.

The fixed-order matching bin-by-bin—as opposed to phase-space point by phase-space point employed with HEJ—does not ensure the same value for the integrated cross section. The effect of the matching will depend on the binning width etc. The size of the variation in the cross sections from the various distributions is one measure of the residual room for improvement in the matching.

The integrated cross sections obtained from various distributions using the method of differential $K$-factors are listed in table 4.1. There is found to be very little variation in the integrated cross section of just 0.1 fb, well within the scale variation on the NLO-rescaled cross section of $6.2^{+1.1}_{-1.2}$ fb, and an overall $K$-factor of 1.6. After

applying VBF-cuts, the overall NLO $K$-factor increases to 2.2, and the NLO-rescaled cross section is found to be $0.53^{+0.15}_{-0.13}$ fb.

Table 4.1 also contains the results for HEJ, which were multiplied by the ratio between HEJ at all orders and NLO both with infinite top-quark mass. Thus the matching is different to that applied at lowest order and normalises to the NLO cross section in the infinite top-quark mass limit. This rescaling is similar to the one used in section 3.2.4. The inclusive cross section for HEJ matched as described is found to be $5.7^{+1.0}_{-1.1}$ fb, slightly lower than the LO prediction for finite top-quark mass multiplied by the NLO $K$-factor from the infinite top-quark mass.

While the inclusive cross sections are similar at NLO and in HEJ, the distributions differ significantly. As is evident from fig. 4.9, the differential distribution from HEJ is harder in $p_{H\perp}$ compared to the scaled LO-result, while the spectrum is decreasing significantly faster for both $m_{12}$ and $\Delta y_{12}$. This means that even though the total cross section for HEJ is matched to NLO (in the infinite top-mass limit) with a scale-dependent $K$-factor of $1.4^{+0.4}_{-0.4}$, within the VBF-cuts the cross section of $0.23^{+0.02}_{-0.03}$ fb happens to be closer, but with a reduced scale dependence, to the LO cross section of $0.24^{+0.12}_{-0.08}$. It is just a numerical coincidence of the cuts applied that the cross sections agree. As seen already in the discussion of the NLO corrections, the perturbative corrections are large in the VBF region. There is no reason to believe the perturbative series has converged already at NLO.

### 4.3.2 Effects of the finite top mass

The impact of the full top-quark mass-dependence on the Born-level calculation for $pp \to H + 2j$ was already investigated in fig. 4.6. While the effect on the integrated cross section is very small, the effect on the differential distribution in $p_{H\perp}$ is enormous. The infinite top-mass approximation undershoots the full-top-mass result by 5% for $p_{H\perp}$ up to 200 GeV and then increasingly overshoots for increasing transverse momentum, reaching 40% discrepancy already at $p_{H\perp} = 340$ GeV. Similarly, for the invariant mass between the two hardest jets, the distribution for the infinite top-mass starts off undershooting the full dependency by 5%, crossing at $m_{12} = 150$ GeV and overshooting by 16% at $m_{12} = 1$ TeV. The error due to the infinite top-quark mass approximation is very small and uniform in the rapidity distribution of the Higgs boson.

We now turn our attention to the impact of both the finite top-quark and bottom-quark mass on HEJ. First, we list in table 4.2 the cross section with inclusive and the VBF cuts for infinite top-quark mass and finite top-quark mass for fixed order

| | Fixed Order | | HEJ | |
|---|---|---|---|---|
| | Inclusive $H + 2j$ | VBF cuts | Inclusive $H + 2j$ | VBF cuts |
| $m_t \to \infty$ | $6.2^{+1.1}_{-1.2}$ fb | $0.54^{+0.16}_{-0.12}$ fb | $6.2^{+1.1}_{-1.2}$ fb | $0.26^{+0.02}_{-0.04}$ fb |
| $m_t = 163\,\text{GeV}$ | $6.2^{+1.1}_{-1.2}$ fb | $0.53^{+0.15}_{-0.13}$ fb | $5.7^{+1.0}_{-1.1}$ fb | $0.23^{+0.02}_{-0.03}$ fb |
| $m_t = 163\,\text{GeV}$<br>$m_b = 2.8\,\text{GeV}$ | - | - | $5.7^{+1.0}_{-1.1}$ fb | $0.23^{+0.02}_{-0.03}$ fb |

Table 4.2: Cross sections obtained in fixed-order perturbation theory (either full NLO using infinite top-quark mass or LO scaled bin-by-bin with the $K$-factor obtained in the infinite top-quark mass limit) and in HEJ for $pp \to H + 2j$ with inclusive and VBF-cuts. See text for further comments.

(LO scaled with NLO in the limit of infinite top-quark mass). For HEJ we also list the results using both finite masses for the top and bottom quark.

The finite top-quark mass has a much larger impact on HEJ than at fixed order, which might at first seem surprising, since the HEJ results are matched to fixed order. The larger impact of the top-mass effects are therefore not a consequence of the approximations in HEJ. Instead, as is evident in the distributions of fig. 4.9, the higher-order corrections of HEJ emphasise the distribution at larger $p_{H\perp}$, where the corrections from the finite quark-mass are large. Therefore, the top-quark mass corrections in HEJ amount to a 9% reduction within the inclusive and 11% within the VBF-cuts. We do not observe any effect of the non-zero bottom-mass beyond 1% for any of the observables studied. The impact obviously increases, if the bottom-mass is chosen larger [164].

Figure 4.10 compares the results obtained with HEJ using the three different descriptions of quark masses, namely infinite top-quark mass, finite $m_t$ but $m_b = 0$, and both $m_t$ and $m_b$ finite. Evidently, the effect of the finite $m_b$ is negligibly small and uniform in all the distributions. As seen already in figs. 4.6 and 4.9, the infinite top mass approximation fails for transverse momenta significantly larger than the top-mass—illustrated here by a plot of the distribution in the transverse momentum of the Higgs boson in fig. 4.10a. Similarly, using an infinite top-mass overshoots the cross section with finite top-mass by 20% at an invariant mass between the two hardest jets of 400 GeV, increasing to 40% at 1 TeV. This is relevant for the description of the contribution from the QCD process within the VBF-studies of $pp \to H + 2j$. The corrections from finite quark masses to the distributions in $\Delta y_{12}$ (fig. 4.10c) or $\Delta\phi_{12}$ with additional VBF cuts (fig. 4.10d) reach just 10%.

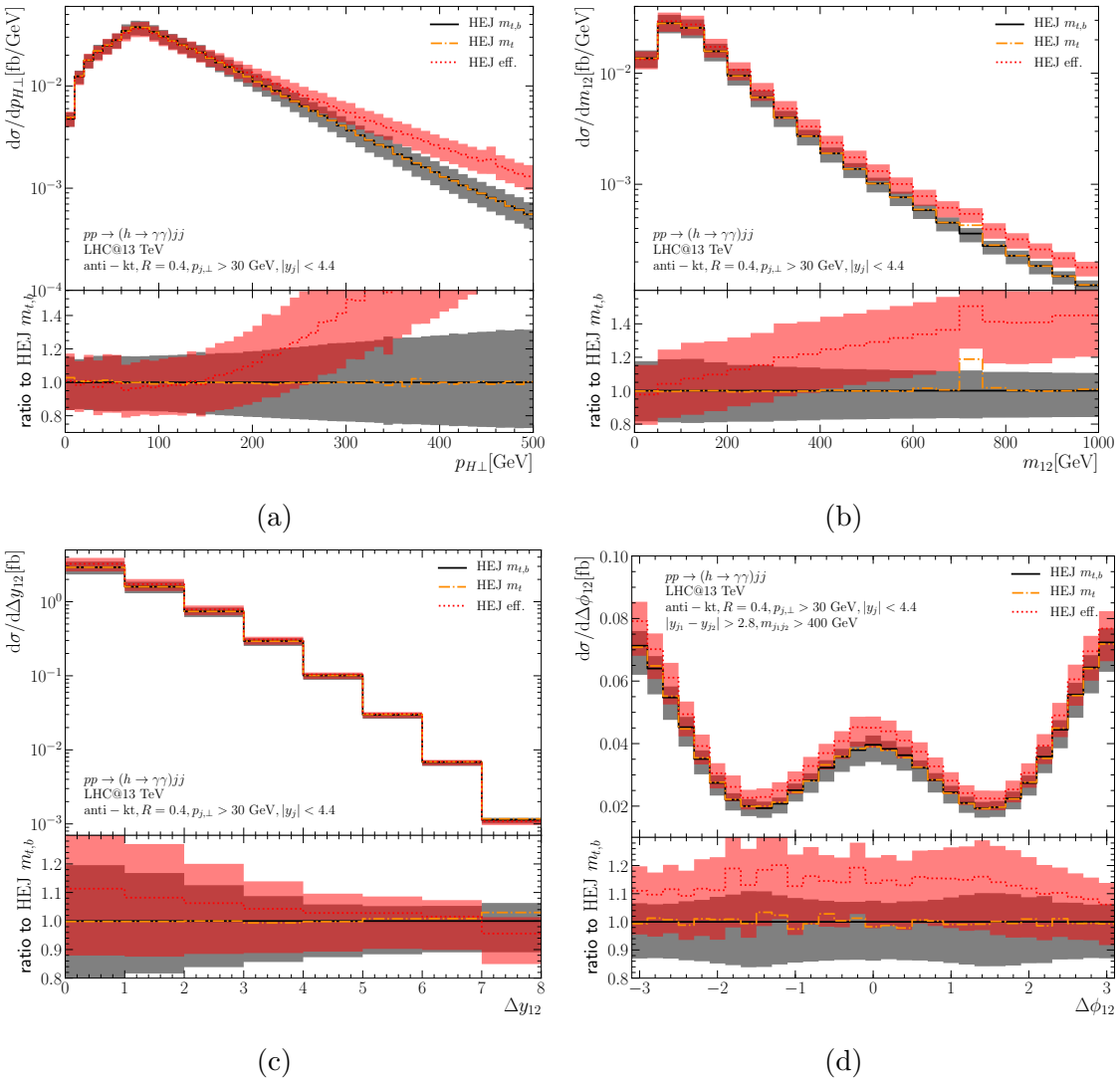



Figure 4.10: Differential cross section of the all-order matched prediction from HEJ with three descriptions of the quark masses: infinite top-quark mass (red, dotted), finite $m_t$ (orange, dot-dashed), and finite $m_t$ and $m_b$ (black/grey, solid). See text for further details.

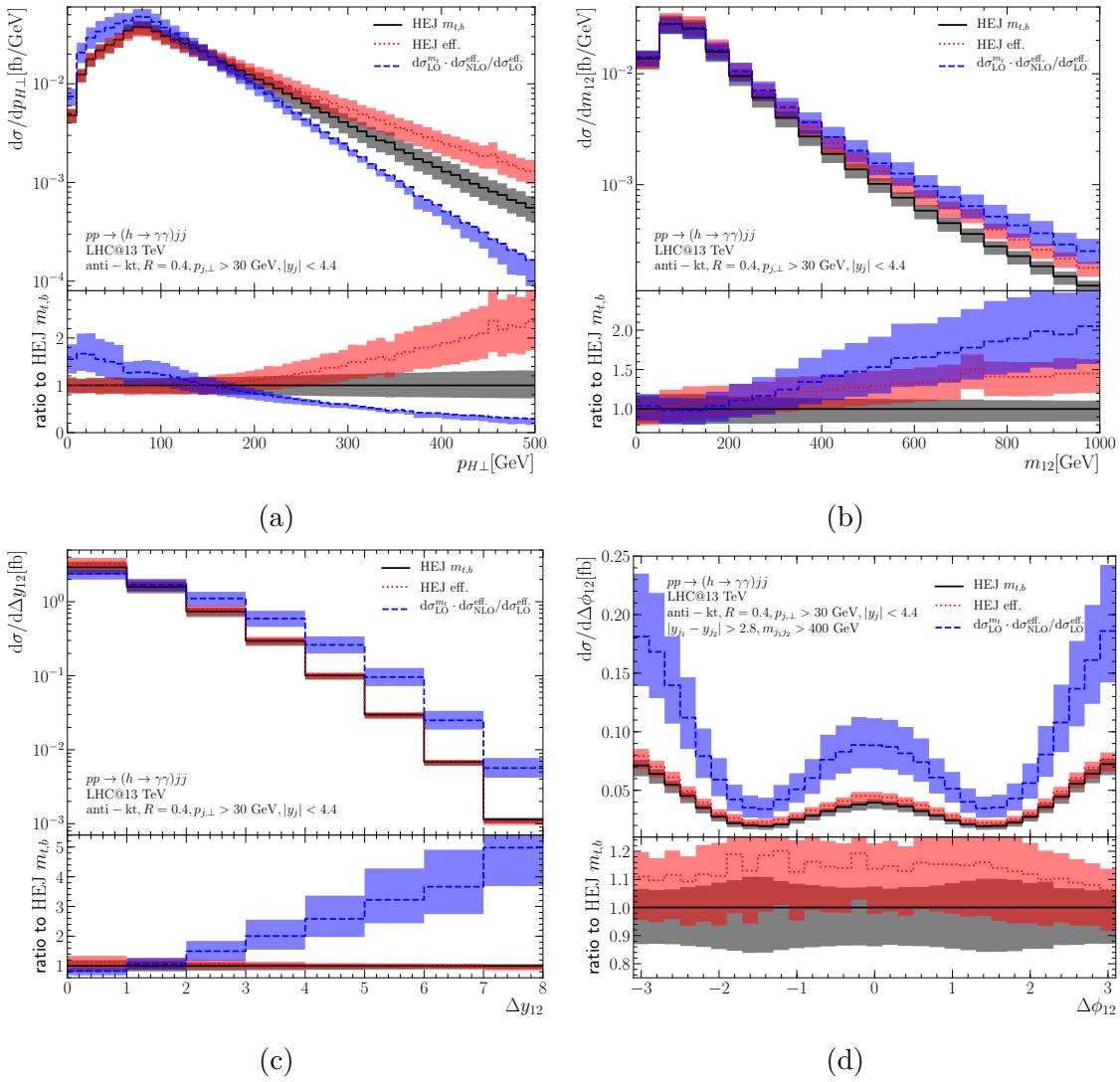



Figure 4.11: The results obtained with HEJ compared with fixed order for various key distributions. See text for further details.

### 4.3.3 Most accurate results

In this section we compare the most accurate predictions obtained using the methods described in this study in HEJ to those obtained at fixed order. We start by comparing the observables already investigated previously; as such, the red and grey bands in fig. 4.11 are identical to those on fig. 4.10, but are here compared to Born-level with finite top-quark mass, rescaled bin-by-bin with the NLO $K$-factor obtained using infinite top-quark mass. We see in fig. 4.11a that the fixed-order result is significantly softer in the transverse momentum of the Higgs boson than that obtained with HEJ. We have already discussed how this leads to a larger impact of the finite quark masses within HEJ than at fixed order.

Figure 4.11b illustrates that the distribution in the invariant mass between the two

hardest jets is increasingly suppressed for increasing $m_{12}$ in HEJ compared to fixed order. While the cross sections are similar for small $m_{12}$, the ratio of fixed order over HEJ reaches 1.5 at $m_{12} \approx 500\,\text{GeV}$. Similarly, as illustrated in fig. 4.11c the results of HEJ are largely suppressed compared to NLO at large $\Delta y_{12}$. The ratio of fixed order to HEJ found here increases linearly finally reaching 5 at $\Delta y_{12} = 8$. As discussed around fig. 4.7, this is due to the absence of a logarithmic suppression of the 3-jet component in the NLO prediction. Although the HEJ cross section is matched to the NLO value, this does not change the differences in the shapes of the distributions and the large $K$-factor at large $m_{12}$ therefore persists.

Figure 4.11d shows the distribution with respect to the azimuthal angle between the two hardest jets, measured relative to the positive rapidity direction, thus exploring the full interval from $-\pi$ to $\pi$. VBF cuts have again been applied in addition to the general cuts. These require a significant invariant mass and rapidity separation of the hardest two jets and hence the suppression in figs. 4.11b and 4.11c translates into a large difference (around a factor of 2) in the cross section between the HEJ and fixed order predictions (as also seen earlier in table 4.2). The distinctive shape which arises as a consequence of the CP structure of the $ggH$ vertex [87, 157, 178] is seen in all the predictions.

We present in figs. C.1 to C.3 the results for the alternative central scale choice of $H_T/2$. The main conclusions of the plots are unchanged; the impact of the higher-order corrections in HEJ lead to a harder distribution in $p_{H\perp}$, which enhances the finite quark mass and loop propagator effects. This in turn leads to a suppression of the prediction at large $m_{12}$, and the predicted impact of a VBF cut is more severe in the all-order calculations of HEJ than that seen in fixed-order predictions.

### 4.3.4 Central jet veto

To end this section we present results for an alternative to a traditional jet veto [87, 114, 143, 157, 178]. We begin by defining two tagging jets $t_1$ and $t_2$: firstly as the hardest two jets in the event $t_{1,2} = j_{1,2}$ and secondly as the most forward/backward jets, $t_{1,2} = j_{f,b}$. We may then construct $y_0 = (y_{t_1} + y_{t_2})/2$ for each event. The event will then be vetoed if it contains a further jet with transverse momentum above 30 GeV in-between the two tagging jets which satisfies $|y_j - y_0| < y_c$ [85]. This procedure applies more focused region in rapidity than a traditional jet veto, which vetos all jets in the gap. Thus the same level of suppression can be obtained with a higher—and therefore perturbatively safer—transverse momentum cut. As discussed in ref. 85 this type of jet veto has, for $y_c$ up to 1.5, very little impact on the VBF process itself, and is an efficient tool in distinguishing the contribution from

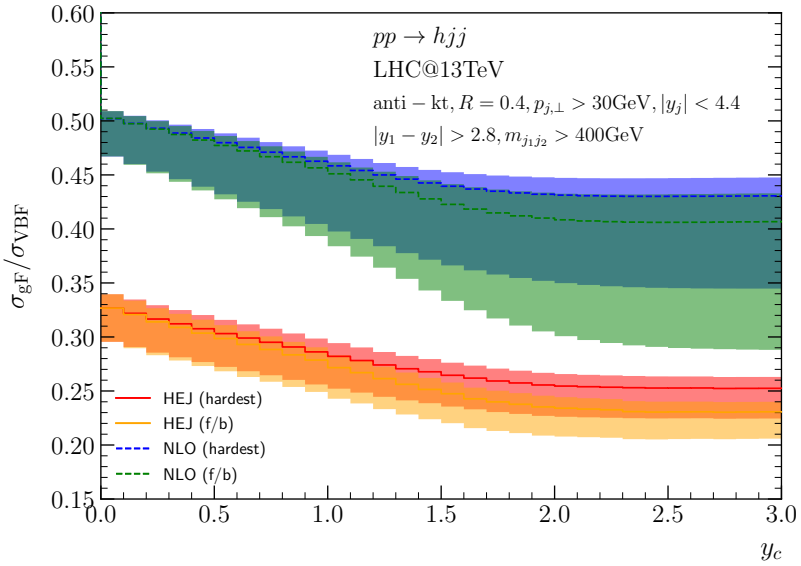


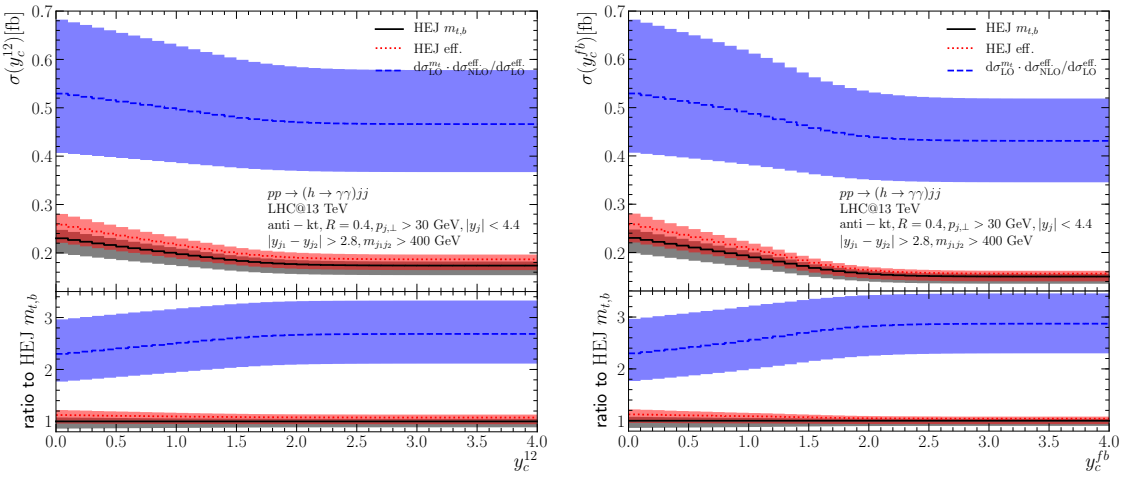

Figure 4.12: Ratio of gluon to weak-boson fusion production for different central jet veto cuts $y_c$. For red and blue the cut is applied to the hardest, and for yellow and green to the most forward and backwards jets. From ref. 85.

the two processes for $pp \to H + 2j$ (see fig. 4.12[5]). This is a direct consequence of the average number of jets over $\Delta y$ we discussed before (cf. figs. 3.14a and 3.14b); extra emissions in VBF are not enhanced with increasing rapidity separation, and therefore mostly produced close in rapidity to the Born-level jets.

In fig. 4.13a, we choose the two hardest jets as the tagging jets, while in fig. 4.13b the tagging jets are the most forward/backward jets. In both cases, the cross section had reached a plateau by about $y_c = 2$. The difference between the two choices is relatively small but the cross section for a given value of $y_c$ is lower for the forward/backward choice for the tagging jets than for the hardest jet choice. We saw that the VBF cuts themselves have a relatively larger impact on the cross sections of HEJ than fixed order, because of the faster reduction in the cross section with $m_{12}$ and $\Delta y_{12}$. Figure 4.13 shows that a further cut on jet activity will have a yet larger effect on HEJ compared to fixed order. This is all expected since the fixed-order results fail to reproduce the rise in jet activity with increasing rapidity separation, which is observed in both data and HEJ [94, 95].

For both selections, the dependence of the finite top-mass on the $y_c$ cut is mild. The main difference comes from the difference cross section after VBF cuts. This is not surprising since the radiation pattern of the extra emission is mostly independent of

[5]The scale choose in ref. 85, and consequently also fig. 4.12, was $\mu_r = \mu_f = H_T/2$. Thus the cross sections in figs. 4.12 and 4.13 are different.




(a) (b)

Figure 4.13: Cross section from HEJ and at NLO as a function of a jet veto, $y_c$, defined in the text. In (a) the tagging jets are the two hardest jets; in (b) the tagging jets are the most forward/backward jets.

the top loop—which is exactly why an inclusion of massive quark-loops in HEJ is possible for any multiplicity.

## 4.4 Summary

We have calculated the gluon-fusion contribution to $H + 2j$ including both

- leading logarithmic corrections in $\hat{s}/p_\perp^2$ to all orders in $\alpha_s$, and
- full dependence on top- and bottom-masses, including the loop-propagator kinematic effects absent in the $m_t \to \infty$ limit.

The components necessary for implementing the full quark-mass dependence within the all-order resummation scheme of High Energy Jets were calculated, such that both the quark mass and the systematic logarithmic corrections within the VBF cuts could be investigated. This goes far beyond the current state-of-the-art fixed order predictions. The results thus obtained have been compared to the fixed-order full top-mass-dependent prediction at Born-level, but rescaled bin-by-bin with the NLO $K$-factor obtained in the limit of $m_t \to \infty$.

While at fixed order the cross section with finite $m_t$ differ very little from those obtained for $m_t \to \infty$, we find a much larger reduction of 9% on the inclusive cross section in HEJ; this is because the transverse momentum of the Higgs boson is found

to be harder with HEJ than at LO, and that the finite mass corrections are larger at large transverse scales. For the first time, our calculation allows the computation of the interference between the top- and bottom-quark contributions beyond leading order in $H + 2j$. We find that the interference is extremely small for the running values of $m_b(m_H)$ and $m_t(m_H)$, less than a percent for all observables.

Further, for a scale choice of $\mu_r = \mu_f = \max(m_H, m_{12})$ the NLO $K$-factor increases systematically for both $m_{12}$, $\Delta y_{12}$ and most dramatically for $\Delta y_{fb}$ (see fig. 4.7). With a scale choice of $\mu_r = \mu_f = H_T/2$ however, the $K$-factor *decreases* with increasing $m_{12}$ or $\Delta y_{12}$. The balance between the large negative virtual corrections and the positive real corrections are clearly scale dependent. The large corrections illustrate a serious perturbative instability of the fixed-order expansion within the VBF-cuts. This instability is specifically addressed by HEJ.

At large $\Delta y_{12}$ and $m_{12}$, the all-order predictions from HEJ are systematically suppressed compared to fixed order. The discussion of scale choice is independent of the discussion of the behaviour at large $m_{12}$ and so is the conclusion that a resummation of the leading terms at large $m_{12}$ leads to a reduction of the cross section within the VBF cuts. Our results show that the gluon-fusion contamination in VBF studies is less severe than the fixed-order estimate would imply. The finite-mass corrections to HEJ within the VBF-cuts lead to a further 11% suppression compared to the assumption of infinite top-mass.

# Chapter 5

# Conclusion

In this thesis we have shown how high energy logarithms are crucial to model accurately collisions at the LHC. We focused at Higgs boson production together with multiple jets, where the inclusion of cuts to suppress the QCD channel from gluon fusion lead to an enhanced sensitivity to logarithms $\log|\hat{s}/\hat{t}|$. We resummed these high energy logarithms in gluon fusion within the framework of HEJ. To increase the accuracy of the result we developed a new matching method to combine HEJ with LO calculations. The new merging works by starting with a fixed order phase space point and adding multiple resummation events. Compared to the previous merging, we therefore need less fixed-order events to reach the same statistical uncertainty. The expensive generation of fixed-order events is factored out into (standard) Monte Carlo generators that specialise in this task, further decreases the computational costs. With the new setup we have been able to match LO and HEJ for up to Higgs boson plus five jets. In principle merging at even higher multiplicity is possible, but gains in the accuracy are marginal even for sensitive observables, while the computational cost would increase significantly.

To further improve the predictions for Higgs plus dijets, we included finite quark-masses. In HEJ this required the addition of new currents for low multiplicity matrix elements, without affecting the implementation of the resummation per se. The finite top mass was shown to be particularly important for large transverse momenta of the Higgs boson. Since HEJ creates harder Higgs bosons compared to fixed order, the difference between finite and infinite quark-mass are more pronounced in the results produced by HEJ. In particular LO showed an accidental cancellation in the total cross section, where the peak of the $p_{H\perp}$ distribution was undershot and the tail overestimated in the infinite top mass approximation. In HEJ we observed a 9 % and 11 % decrease in the inclusive and VBF selection cross section respectively.

The code used to generate all predictions in this thesis has been publicly released as

version 2 of HEJ (see ref. 4, 5). More processes, like $W$ and $Z$ boson with jets, are in development, and will be made available in the next minor release. Additionally HEJ 2 offers a much more flexible code base, which will utilised further in the future, e.g. through an reimplementation of the HEJ with Pythia merging [99] for arbitrary process. The new merging algorithm should also be generalisable for merging HEJ with NLO predictions, which would improve the perturbative stability and reduce the scale variations.

# Appendix A

# Spinor-helicity formalism

Calculating amplitudes in QCD can become quite involved, with many different terms. The *spinor-helicity formalism* can help to organise such calculations more efficiently [39, 179]. We start by separating the amplitudes into a colour flow and a *colour-stripped amplitude*. As the name suggests, the colour flow gives the colour exchange in the colour space. The colour-stripped amplitude contains all kinematic effects, which can be expressed through the spinor-helicity formalism. Each (Dirac) spinor gets spit according to its helicity into a left- and right-handed (Weyl) spinor respectively

$$|i\rangle = u^+(p_i), \qquad |i] = u^-(p_i), \tag{A.1.1}$$

$$\langle i| = u^{-\dagger}(p_i)\gamma^0, \qquad [i| = u^{+\dagger}(p_i)\gamma^0, \tag{A.1.2}$$

where $p_i$ is the momentum of particle $i$. Similarly gluon $i$ can be expressed as

$$\epsilon_\mu^+ = \frac{1}{\sqrt{2}}\frac{\langle r|\mu|i]}{\langle r\, i\rangle} \qquad \Leftrightarrow \qquad \not{\epsilon}^+ = \sqrt{2}\frac{|i]\,\langle r| + |r\rangle\,[i|}{\langle r\, i\rangle}\,, \tag{A.1.3}$$

$$\epsilon_\mu^- = \frac{1}{\sqrt{2}}\frac{[r|\mu|i\rangle}{[r\, i]} = \epsilon_\mu^{+*} \qquad \Leftrightarrow \qquad \not{\epsilon}^- = \sqrt{2}\frac{|i\rangle\,[r| + |r]\,\langle i|}{[r\, i]}\,, \tag{A.1.4}$$

with $[i|\mu|j\rangle = [i|\,\gamma^\mu\,|j\rangle$. The polarisation of each gluon is fixed by an arbitrary reference momentum $r$ satisfying $r \cdot p_i \neq 0$. The spinor-helicity formalism removes redundancies in calculations, since each contractions can be related back to Mandelstam variables[1]

$$s_{ij} = 2\, p_i \cdot p_j = \langle i\, j\rangle\,[j\, i]\,. \tag{A.1.5}$$

[1] We assume all particles to be massless, i.e. $p_i^2 = 0\ \forall i$.

The product of two spinors with different helicity is antisymmetric

$$\langle i\, j\rangle = -\langle j\, i\rangle\,, \tag{A.1.6}$$

$$[i\, j] = -[j\, i]\,, \tag{A.1.7}$$

For real momenta with the same signed energies we further find $[i\, j] = \langle j\, i\rangle^*$. Contracting only one helicity gives zero

$$\langle i\, j] = [i\, j\rangle = 0. \tag{A.1.8}$$

We can generalise this by adding extra $\gamma$ matrices; each contraction with an even number of $\gamma$ matrices is antisymmetric

$$\langle i|\mu_1 \cdots \mu_{2N}|j\rangle = -\langle j|\mu_{2N} \cdots \mu_1|i\rangle \tag{A.1.9}$$

and with an odd number symmetric

$$\langle i|\mu_1 \cdots \mu_{2N+1}|j] = [j|\mu_{2N+1} \cdots \mu_1|i\rangle\,, \tag{A.1.10}$$

$$\langle i|\mu|j\rangle = [i|\mu|j] = 0. \tag{A.1.11}$$

We further realise that

$$\not{p}_i = |i\rangle\,[i| + |i]\,\langle i|\,. \tag{A.1.12}$$

$$\tag{A.1.13}$$

Hence, each momentum reduce the number of $\gamma$ matrices by one. We are left with four basic cases

$$\langle i|\mu_1 \ldots \mu_I \not{\mu}_{I+1} \ldots \mu_{2N+1}|j] = \begin{cases} \langle i|\mu_1 \ldots \mu_I|r]\,\langle r|\mu_{I+1} \ldots \mu_{2N+1}|j] & I \text{ even} \\ \langle i|\mu_1 \ldots \mu_I|r\rangle\,[r|\mu_{I+1} \ldots \mu_{2N+1}|j] & I \text{ odd} \end{cases}, \tag{A.1.14}$$

$$\langle i|\mu_1 \ldots \mu_I \not{\mu}_{I+1} \ldots \mu_{2N}|j\rangle = \begin{cases} \langle i|\mu_1 \ldots \mu_I|r]\,\langle r|\mu_{I+1} \ldots \mu_{2N}|j\rangle & I \text{ even} \\ \langle i|\mu_1 \ldots \mu_I|r\rangle\,\langle r|\mu_{I+1} \ldots \mu_{2N}|j]^* & I \text{ odd} \end{cases}. \tag{A.1.15}$$

These cases can be automatically be employed to expand amplitudes with any number of $\gamma$ matrices, such that we are left with only terms containing either none or exactly one open index. The terms remaining open indices have to be contracted over different terms via the *Fierz identity*

$$\langle i|\mu|j]\,\langle r|\mu|s] = 2\,\langle i\, r\rangle\,[s\, j]\,. \tag{A.1.16}$$

In total we can express any (tree-level) amplitude to a multiplication of terms like $\langle i\, j\rangle$ and $[i\, j]$, which is essentially a multiplication of complex numbers.

As a simple example we consider the scattering of two quarks with different flavours. All helicity configurations in which only one helicity is different from the rest vanish. We are left with only four amplitudes, and two unique terms

$$\mathcal{M}_{q^-Q^-\to q^-Q^-} = \frac{1}{t}\langle 1|\mu|a]\,\langle 2|\mu|b] \tag{A.1.17}$$

$$= \frac{2}{t}\langle 1\,2\rangle\,[b\,a] = \mathcal{M}^*_{q^+Q^+\to q^+Q^+}\,, \tag{A.1.18}$$

$$\mathcal{M}_{q^+Q^-\to q^+Q^-} = \frac{1}{t}[1|\mu|a\rangle\,\langle 2|\mu|b] \tag{A.1.19}$$

$$= \frac{2}{t}\langle a\,2\rangle\,[b\,1] = \mathcal{M}^*_{q^-Q^+\to q^-Q^+}\,. \tag{A.1.20}$$

In the second and fourth line we used the Fierz identity from eq. (A.1.16). The remaining terms already show the structure of eq. (A.1.5). We can then directly read off the helicity averaged (colour stripped) matrix element

$$\overline{\left|\mathcal{M}_{qQ\to qQ}\right|}^2 = 2\frac{s^2+u^2}{t^2}\,. \tag{A.1.21}$$

For completeness we also repeat the definition of the currents from eq. (2.3.8) here

$$j^-_\mu(p_i,p_j) = \langle i|\mu|j]\,, \tag{A.1.22}$$

$$j^+_\mu(p_i,p_j) = [i|\mu|j\rangle = j^-_\mu(p_j,p_i). \tag{A.1.23}$$

With this the structure of eqs. (A.1.18) and (A.1.20) gets even more obvious

$$\mathcal{M}_{q^-Q^-\to q^-Q^-} = \frac{1}{t}j^-_\mu(p_1,p_a)j^-_\mu(p_2,p_b), \tag{A.1.24}$$

$$\mathcal{M}_{q^+Q^-\to q^+Q^-} = \frac{1}{t}j^+_\mu(p_1,p_a)j^-_\mu(p_2,p_b). \tag{A.1.25}$$

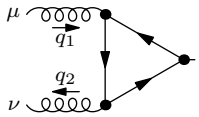
# Appendix B

# HEJ currents with finite quark masses

## B.1 Form factors for the Higgs-boson coupling to gluons

Quoting eq. (4.1.1), the coupling of the Higgs boson to gluons via a virtual quark loop can be written as

$$V_H^{\mu\nu}(q_1,q_2) = \quad = \frac{\alpha_s m^2}{\pi v}\Big[g^{\mu\nu}T_1(q_1,q_2) - q_2^\mu q_1^\nu T_2(q_1,q_2)\Big]. \quad \text{(B.1.1)}$$




The outgoing momentum of the Higgs boson is $p_H = q_1 - q_2$. The form factors $T_1$ and $T_2$ are then given by [109]

$$\begin{aligned} T_1(q_1,q_2) = & -C_0(q_1,q_2)\left[2m^2 + \frac{1}{2}\left(q_1^2+q_2^2-p_H^2\right) + \frac{2q_1^2q_2^2p_H^2}{\lambda}\right] \\ & - \left[B_0(q_2) - B_0(p_H)\right]\frac{q_2^2}{\lambda}\left(q_2^2-q_1^2-p_H^2\right) \\ & - \left[B_0(q_1) - B_0(p_H)\right]\frac{q_1^2}{\lambda}\left(q_1^2-q_2^2-p_H^2\right) - 1, \end{aligned} \quad \text{(B.1.2)}$$

$$
\begin{aligned}
T_2(q_1,q_2) = C_0(q_1,q_2) \Bigg[ & \frac{4m^2}{\lambda}\left(p_H^2 - q_1^2 - q_2^2\right) - 1 - \frac{4q_1^2 q_2^2}{\lambda} \\
& - \frac{12 q_1^2 q_2^2 p_H^2}{\lambda^2}\left(q_1^2 + q_2^2 - p_H^2\right)\Bigg] \\
& - \left[B_0(q_2) - B_0(p_H)\right]\left[\frac{2q_2^2}{\lambda} + \frac{12 q_1^2 q_2^2}{\lambda^2}\left(q_2^2 - q_1^2 + p_H^2\right)\right] \\
& - \left[B_0(q_1) - B_0(p_H)\right]\left[\frac{2q_1^2}{\lambda} + \frac{12 q_1^2 q_2^2}{\lambda^2}\left(q_1^2 - q_2^2 + p_H^2\right)\right] \\
& - \frac{2}{\lambda}\left(q_1^2 + q_2^2 - p_H^2\right),
\end{aligned}
\tag{B.1.3}
$$

where we have used the scalar bubble and triangle integrals

$$
B_0(p) = \int \frac{d^d l}{i\pi^{\frac{d}{2}}} \frac{1}{\left(l^2 - m^2\right)\left((l+p)^2 - m^2\right)}, \tag{B.1.4}
$$

$$
C_0(p,q) = \int \frac{d^d l}{i\pi^{\frac{d}{2}}} \frac{1}{\left(l^2 - m^2\right)\left((l+p)^2 - m^2\right)\left((l+p-q)^2 - m^2\right)}, \tag{B.1.5}
$$

and the Källén function

$$
\lambda = q_1^4 + q_2^4 + p_H^4 - 2q_1^2 q_2^2 - 2q_1^2 p_H^2 - 2q_2^2 p_H^2. \tag{B.1.6}
$$

The relation to the form factors $A_1, A_2$ given in ref. 109 is

$$
A_1(q_1,q_2) = \frac{i}{16\pi^2} T_2(-q_1,q_2), \tag{B.1.7}
$$

$$
A_2(q_1,q_2) = -\frac{i}{16\pi^2} T_1(-q_1,q_2). \tag{B.1.8}
$$

In the infinite top-mass limit

$$
\lim_{m\to\infty} m^2 T_1(q_1,q_2) = \frac{1}{3} q_1 \cdot q_2\,, \tag{B.1.9}
$$

$$
\lim_{m\to\infty} m^2 T_2(q_1,q_2) = \frac{1}{3}\,. \tag{B.1.10}
$$

Thus eq. (B.1.1) converges to eq. (2.3.17).

## B.2 Effective current for peripheral Higgs-boson emission

We describe the emission of a peripheral Higgs boson close to a scattering gluon with an effective current. In the following we consider a light-cone decomposition of the gluon momenta, i.e. $p^{\pm} = E \pm p_z$ and $p_{\perp} = p_x + ip_y$. The incoming gluon momentum $p_a$ defines the $-$ direction, so that $p_a^+ = p_{a\perp} = 0$. The outgoing momenta are $p_1$ for the gluon and $p_H$ for the Higgs boson. We choose the following polarisation vectors:

$$\epsilon_{\mu}^{\pm}(p_a) = \frac{j_{\mu}^{\pm}(p_1, p_a)}{\sqrt{2}\bar{u}^{\pm}(p_a)u^{\mp}(p_1)}\,, \quad \epsilon_{\mu}^{\pm,*}(p_1) = -\frac{j_{\mu}^{\pm}(p_1, p_a)}{\sqrt{2}\bar{u}^{\mp}(p_1)u^{\pm}(p_a)}\,. \tag{B.2.1}$$

Following ref. 86, we introduce effective polarisation vectors to describe the contraction with the Higgs-boson production vertex eq. (4.1.1):

$$\epsilon_{H,\mu}(p_a) = \frac{T_2(p_a, p_a - p_H)}{(p_a - p_H)^2}\Big[p_a \cdot p_H \epsilon_{\mu}(p_a) - p_H \cdot \epsilon(p_a) p_{a,\mu}\Big], \tag{B.2.2}$$

$$\epsilon_{H,\mu}^{*}(p_1) = -\frac{T_2(p_1 + p_H, p_1)}{(p_1 + p_H)^2}\Big[p_1 \cdot p_H \epsilon_{\mu}^{*}(p_1) - p_H \cdot \epsilon^{*}(p_1) p_{1,\mu}\Big], \tag{B.2.3}$$

We also employ the short-hand notation from appendix A and

$$[i|H|j\rangle = j_{\mu}^{+}(p_i, p_j) p_H^{\mu}\,. \tag{B.2.4}$$

Without loss of generality, we consider only the case where the incoming gluon has positive helicity. The remaining helicity configurations can be obtained through parity transformation.

Labelling the effective current by the helicities of the gluons we obtain for the same-helicity case

$$\begin{aligned}
j_{H,\mu}^{++}(p_1, p_a, p_H) = \frac{m^2}{\pi v}\Bigg[ & \\
& - \sqrt{\frac{2p_1^-}{p_a^-}} \frac{p_{1\perp}^*}{|p_{1\perp}|} \frac{t_2}{[a\,1]} \epsilon_{H,\mu}^{+,*}(p_1) + \sqrt{\frac{2p_a^-}{p_1^-}} \frac{p_{1\perp}^*}{|p_{1\perp}|} \frac{t_2}{\langle 1\,a\rangle} \epsilon_{H,\mu}^{+}(p_a) \\
& + [1|H|a\rangle \left( \frac{\sqrt{2}}{\langle 1\,a\rangle} \epsilon_{H,\mu}^{+}(p_a) + \frac{\sqrt{2}}{[a\,1]} \epsilon_{H,\mu}^{+,*}(p_1) - \frac{\langle 1\,a\rangle\, T_2(p_a, p_a - p_H)}{\sqrt{2}(p_a - p_H)^2} \epsilon_{\mu}^{+,*}(p_1) \right. \\
& \qquad \left. - \frac{[a\,1]\, T_2(p_1 + p_H, p_1)}{\sqrt{2}(p_1 + p_H)^2} \epsilon_{\mu}^{+}(p_a) - \frac{RH_4}{\sqrt{2}\,[a\,1]} \epsilon_{\mu}^{+,*}(p_1) + \frac{RH_5}{\sqrt{2}\,\langle 1\,a\rangle} \epsilon_{\mu}^{+}(p_a) \right) \\
& - \frac{[1|H|a\rangle^2}{2t_1} (p_{a,\mu} RH_{10} - p_{1,\mu} RH_{12}) \Bigg]
\end{aligned} \tag{B.2.5}$$

with $t_1 = (p_a - p_1)^2$, $t_2 = (p_a - p_1 - p_H)^2$ and $R = 8\pi^2$. The form factors $H_i$ are

given in ref. 109.

The current with a flip in the gluon helicity reads

$$
\begin{aligned}
j^{+-}_{H,\mu}(p_1,p_a,p_H) = \frac{m^2}{\pi v}\Bigg[ & \\
& - \sqrt{\frac{2p_1^-}{p_a^-}}\frac{p_{1\perp}^*}{|p_{1\perp}|}\frac{t_2}{[a\,1]}\epsilon^{-,*}_{H,\mu}(p_1) + \sqrt{\frac{2p_a^-}{p_1^-}}\frac{p_{1\perp}}{|p_{1\perp}|}\frac{t_2}{[a\,1]}\epsilon^{+}_{H,\mu}(p_a) \\
& + [1|H|a\rangle\left(\frac{\sqrt{2}}{[a\,1]}\epsilon^{-,*}_{H,\mu}(p_1) - \frac{\langle 1\,a\rangle\, T_2(p_a,p_a-p_H)}{\sqrt{2}(p_a-p_H)^2}\epsilon^{-,*}_{\mu}(p_1) - \frac{RH_4}{\sqrt{2}\,[a\,1]}\epsilon^{-,*}_{\mu}(p_1)\right) \\
& + [a|H|1\rangle\left(\frac{\sqrt{2}}{[a\,1]}\epsilon^{+}_{H,\mu}(p_a) - \frac{\langle 1\,a\rangle\, T_2(p_1+p_H,p_1)}{\sqrt{2}(p_1+p_H)^2}\epsilon^{+}_{\mu}(p_a) + \frac{RH_5}{\sqrt{2}\,[a\,1]}\epsilon^{+}_{\mu}(p_a)\right) \\
& - \frac{[1|H|a\rangle\,[a|H|1\rangle}{2\,[a\,1]^2}(p_{a,\mu}RH_{10} - p_{1,\mu}RH_{12}) \\
& + \frac{\langle 1\,a\rangle}{[a\,1]}\Bigg(RH_1 p_{1,\mu} - RH_2 p_{a,\mu} + 2p_1\cdot p_H\frac{T_2(p_1+p_H,p_1)}{(p_1+p_H)^2}p_{a,\mu} \\
& \quad - 2p_a\cdot p_H\frac{T_2(p_a,p_a-p_H)}{(p_a-p_H)^2}p_{1,\mu} + T_1(p_a-p_1,p_a-p_1-p_H)\frac{(p_1+p_a)_\mu}{t_1} \\
& \quad - \frac{(p_1+p_a)\cdot p_H}{t_1}T_2(p_a-p_1,p_a-p_1-p_H)(p_1-p_a)_\mu\Bigg)\Bigg].
\end{aligned}
\tag{B.2.6}
$$

If we instead choose the gluon momentum in the + direction, so that $p_a^- = p_{a\perp} = 0$, the corresponding currents are obtained by replacing $p_1^- \to p_1^+, p_a^- \to p_a^+, \frac{p_{1\perp}}{|p_{1\perp}|} \to -1$ in the second line of eq. (B.2.5) and eq. (B.2.6).

## B.3 The current for a single unordered gluon emission

In section 4.1.4, we use an effective current, $j_\mu^{\text{uno cd}}(p_2, p_1, p_a)$, to describe the emission of an unordered gluon (one additional gluon outside in rapidity of an FKL configuration). The current for $q(p_a) \to g(p_1)q(p_2)g^*(\tilde{q}_2)$ was derived in ref. 98 to be

$$j^{\text{uno}\,\mu\ cd}(p_2, p_1, p_a) = i\varepsilon_{1\nu}\left(T^c_{2i}T^d_{ia}\ (U_1^{\mu\nu} - L^{\mu\nu}) + T^d_{2i}T^c_{ia}\ (U_2^{\mu\nu} + L^{\mu\nu})\right). \tag{B.3.1}$$

where

$$U_1^{\mu\nu} = \frac{1}{s_{21}}\left(j_{21}^\nu j_{1a}^\mu + 2p_2^\nu j_{2a}^\mu\right), \qquad U_2^{\mu\nu} = \frac{1}{t_{a1}}\left(2j_{2a}^\mu p_a^\nu - j_{21}^\mu j_{1a}^\nu\right), \tag{B.3.2}$$

$$L^{\mu\nu} = \frac{1}{t_{a2}}\left(-2p_1^\mu j_{2a}^\nu + 2p_1.j_{2a}g^{\mu\nu} + (\tilde{q}_1 + \tilde{q}_2)^\nu j_{2a}^\mu + \frac{t_{b2}}{2}j_{2a}^\mu\left(\frac{p_2^\nu}{p_1.p_2} + \frac{p_b^\nu}{p_1.p_b}\right)\right), \tag{B.3.3}$$

$\tilde{q}_1 = p_a - p_1$ and $\tilde{q}_2 = \tilde{q}_1 - p_2$. This differs from our other currents as there is no longer a single overall colour factor, and hence colour factors (with free indices $c$ and $d$) must be included. Upon contracting with another current squaring, this leads to terms with different colour factors. For example, for $q(p_a)Q(p_b) \to g(p_1)q(p_2)Q(p_3)$, we find [98]

$$\begin{aligned}\left|\overline{\mathcal{M}^{\text{HEJ}}_{\text{tree }qQ\to gqQ}}\right|^2 &= -\frac{g_s^6}{16t_{b3}^2}\sum_{h_a,h_1,h_b,h_2} C_F\Bigg[2\text{Re}\left(\left[j_{3b\,\mu}(L^{\mu\nu} - U_1^{\mu\nu})\right]\left[j_{3b\,\rho}(L^\rho{}_\nu + U_2^\rho{}_\nu)\right]^*\right) \\ &\qquad + 2\frac{C_F}{C_A}\left|j_{3b_\mu}(U_1^{\mu\nu} + U_2^{\mu\nu})\right|^2\Bigg] \\ &\equiv -\frac{g_s^6}{16t_{b3}^2}C_F\left\|\mathcal{S}^{\text{uno}}_{f_1f_2\to gf_1f_2}\right\|^2.\end{aligned} \tag{B.3.4}$$

The factor we require in eq. (4.1.17) is therefore given by

$$\begin{aligned}\left\|\mathcal{S}^{\text{uno}}_{qf_2\to gqHf_2}\right\|^2 = \sum_{h_a,h_1,h_b,h_2}\Bigg[&2\text{Re}\left(\left[J_\mu(L^{\mu\nu} - U_1^{\mu\nu})\right]\left[J_\rho(L^\rho{}_\nu + U_2{}^\rho{}_\nu)\right]^*\right) \\ &+ 2\frac{C_F}{C_A}\left|J_\mu(U_1^{\mu\nu} + U_2^{\mu\nu})\right|^2\Bigg],\end{aligned} \tag{B.3.5}$$

where we use the shorthand $J^\mu = V_H^{\mu\nu}(q_1, q_2)j_\nu(p_3, p_b)$.

# Appendix C

# Finite quark mass results with $\mu_r = \mu_f = H_T/2$

In this appendix we study the effect of using a central scale of $\mu_r = \mu_f = H_T/2$ instead of $\mu_r = \mu_f = \max(m_H, m_{12})$ used in chapter 4. In table C.1 we present the cross section results for a central scale choice of $\mu_r = \mu_f = H_T/2$. These correspond to the results in table 4.2 in section 4.3.2. We continue in figs. C.1 to C.3 by repeating the comparisons of figs. 4.9 to 4.11. While there are variations in numerical values, we find that the conclusions of the impact of the higher-order corrections in HEJ and of the finite quark mass and loop propagator effects are unchanged.

The results obtained at NLO for the two central scale choices $\mu_r = \mu_f = \max(m_H, m_{12})$ and $\mu_r = \mu_f = H_T/2$ are compared in fig. C.4. It is noteworthy that the difference in the results in fig. C.4d for the cross section within the VBF-cuts is similar to the difference between the results of NLO and HEJ obtained with the same scale.

Finally, fig. C.5 compares the results obtained for HEJ with the two central scale choices. The differences in the results for the distributions are significantly larger

| | Fixed Order | | HEJ | |
|---|---|---|---|---|
| | Inclusive $H + 2j$ | VBF cuts | Inclusive $H + 2j$ | VBF cuts |
| $m_t \to \infty$ | $6.4^{+0.3}_{-0.9}$ fb | $0.82^{+0.02}_{-0.11}$ fb | $6.4^{+0.3}_{-0.9}$ fb | $0.56^{+0.04}_{-0.09}$ fb |
| $m_t = 163\,\text{GeV}$ | $6.6^{+0.3}_{-1.0}$ fb | $0.82^{+0.02}_{-0.11}$ fb | $6.2^{+0.3}_{-0.9}$ fb | $0.51^{+0.03}_{-0.08}$ fb |
| $m_t = 163\,\text{GeV}$<br>$m_b = 2.8\,\text{GeV}$ | - | - | $6.2^{+0.3}_{-0.9}$ fb | $0.52^{+0.03}_{-0.08}$ fb |

Table C.1: Total cross section predictions for the central scale choice $\mu_r = \mu_f = H_T/2$ and different values of the heavy-quark masses. See table 4.2 for the corresponding predictions with $\mu_r = \mu_f = \max(m_H, m_{12})$.

than indicated by the scale variation. This is not surprising, since the leading logarithmic behaviour at large $m_{12}$ is unrelated to $\beta_0$-terms from the running of the coupling. As stated earlier, comparisons with data for other processes can determine which of these scale choices obtains the best description. The discussion of scale choice is independent of the discussion of the behaviour at large $m_{12}$, and so is the conclusion that a resummation of the leading terms at large $m_{12}$ leads to a reduction of the cross section within the VBF cuts.

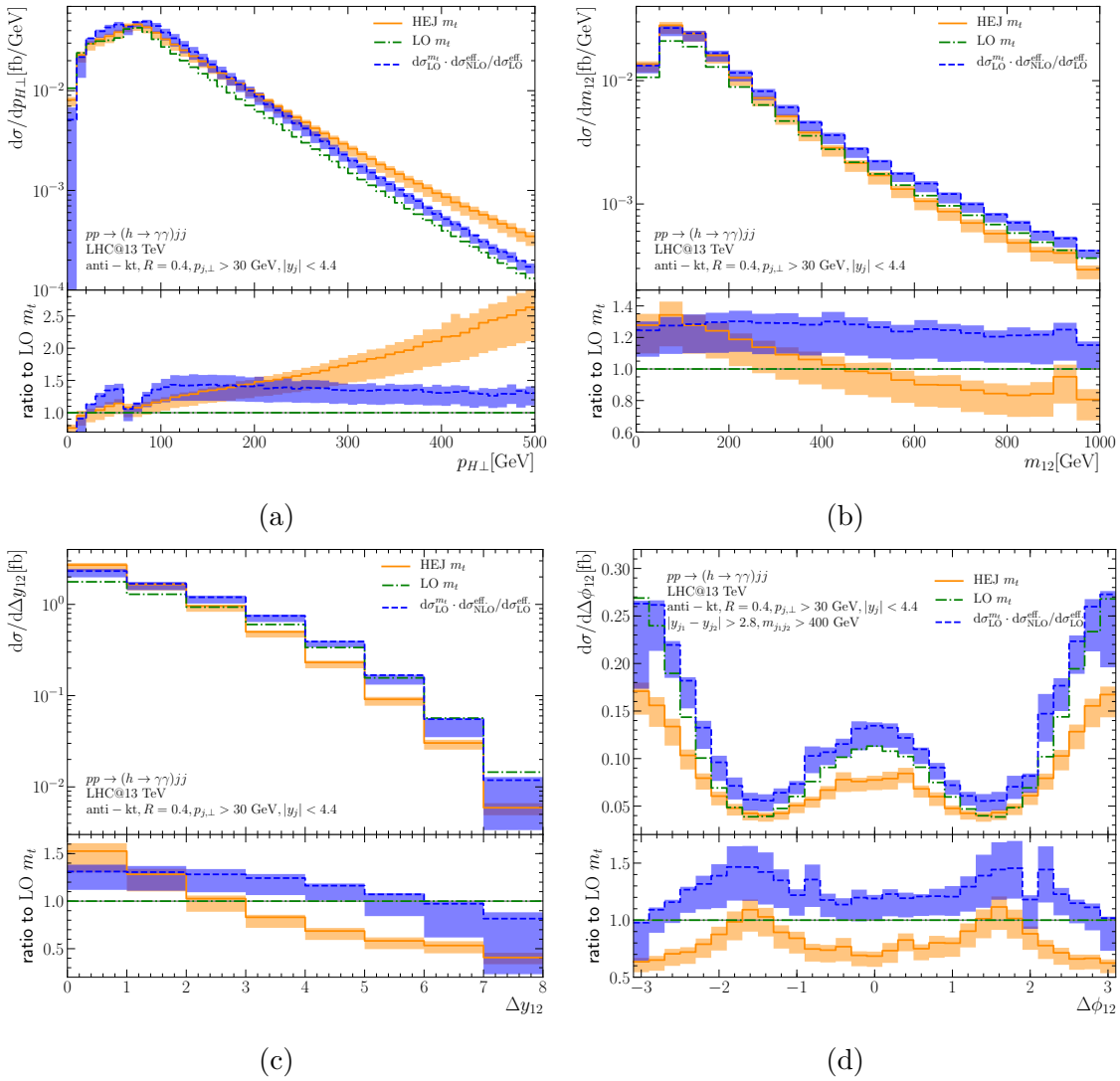



Figure C.1: Predictions for various distributions obtained with HEJ, pure leading order, and leading order rescaled with differential $K$ factors for the central scale choice $\mu_r = \mu_f = H_T/2$. See fig. 4.9 for the corresponding plots with $\mu_r = \mu_f = \max(m_H, m_{12})$.

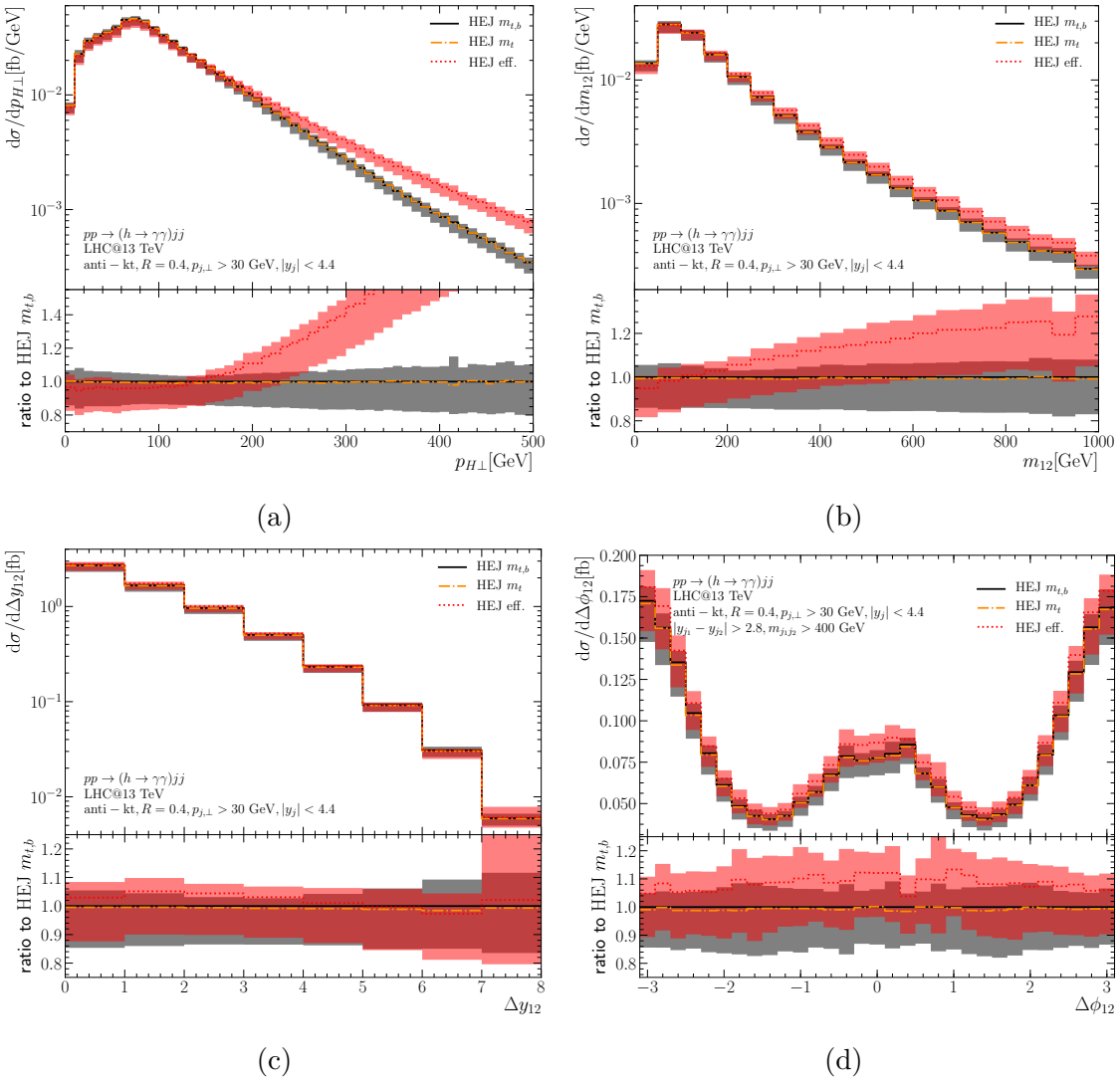



Figure C.2: HEJ predictions for various distributions and different choices for the heavy-quark mass with the central scale choice $\mu_r = \mu_f = H_T/2$. See fig. 4.10 for the corresponding plots with $\mu_r = \mu_f = \max(m_H, m_{12})$.

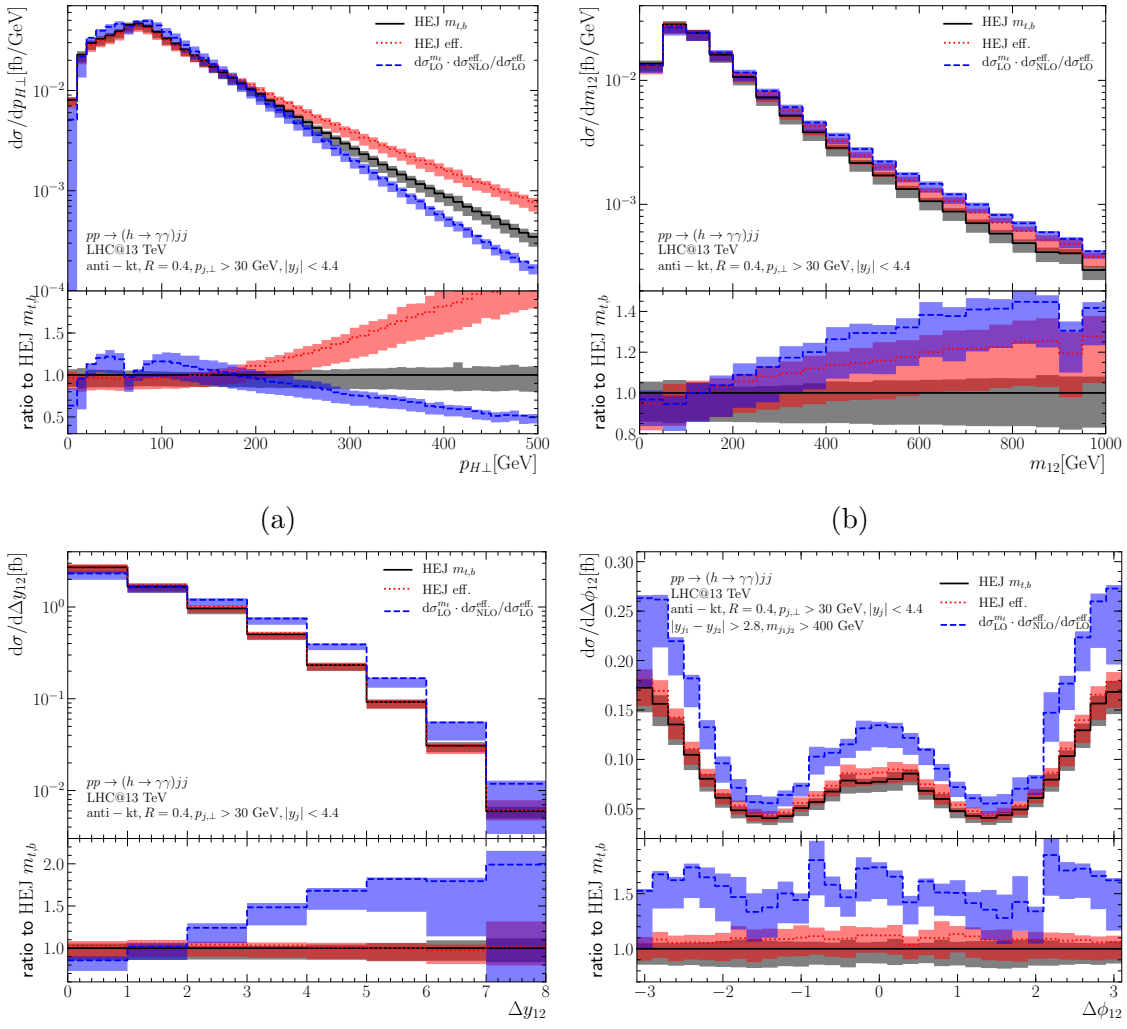



Figure C.3: Comparison between HEJ and the rescaled leading-order prediction for various distributions with the central scale choice $\mu_r = \mu_f = H_T/2$. See fig. 4.11 for the corresponding plots with $\mu_r = \mu_f = \max(m_H, m_{12})$

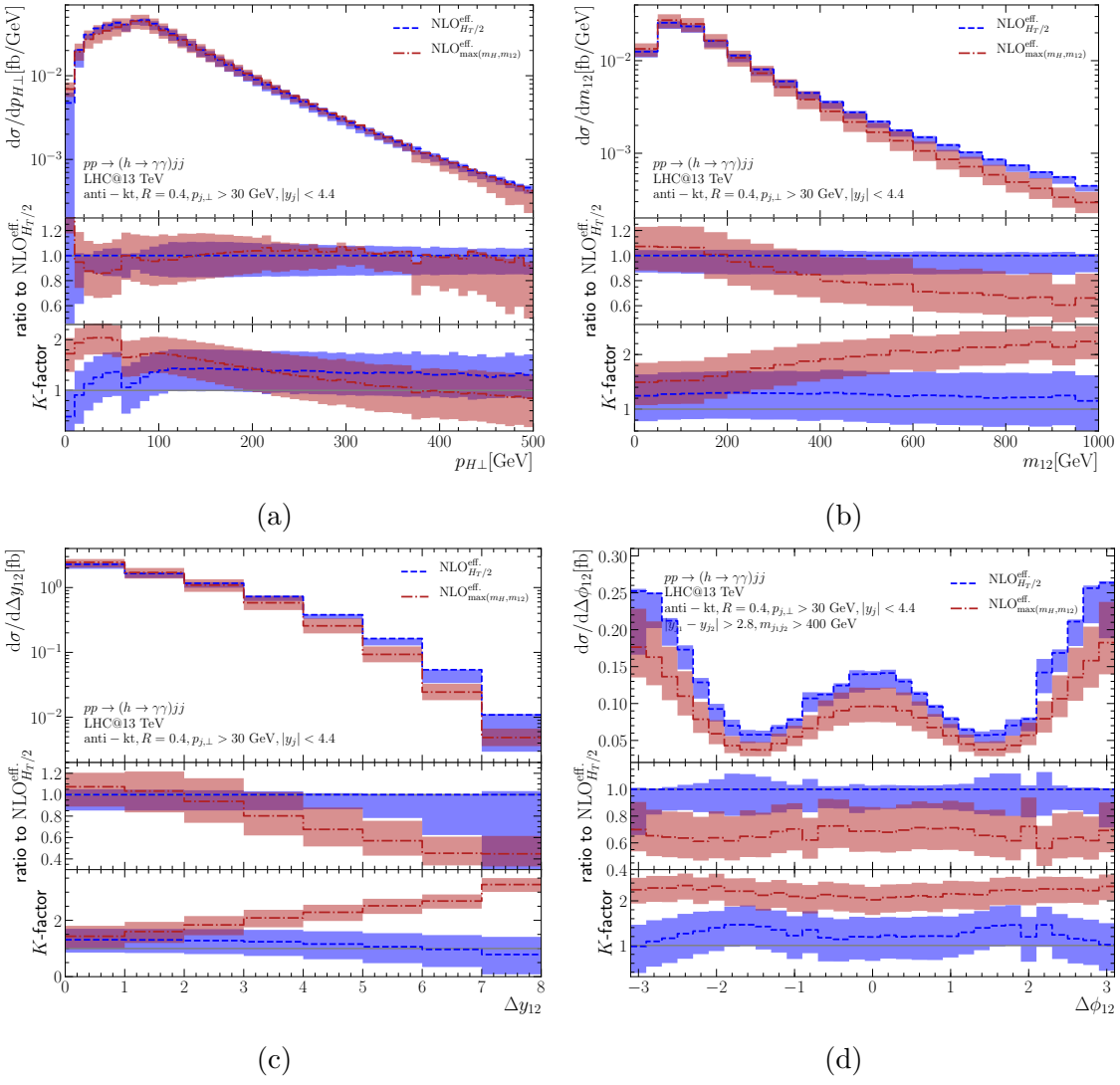



Figure C.4: Comparison between pure NLO results with central scale choices $\mu_r = \mu_f = H_T/2$ and $\mu_r = \mu_f = \max(m_H, m_{12})$

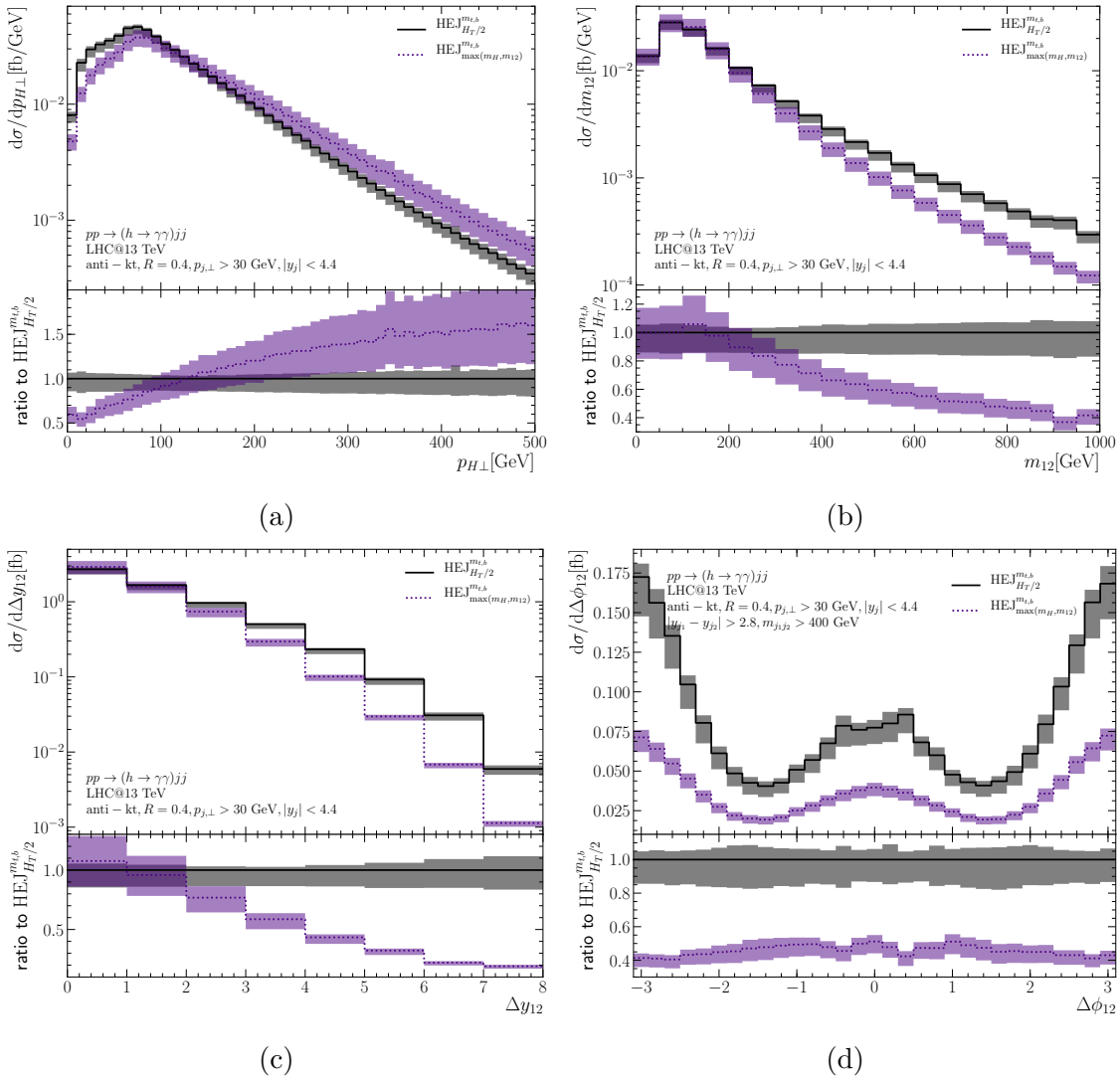



Figure C.5: Comparison between HEJ results with central scale choices $\mu_r = \mu_f = H_T/2$ and $\mu_r = \mu_f = \max(m_H, m_{12})$

# Bibliography


[1] J.R. Andersen, T. Hapola, M. Heil, A. Maier and J.M. Smillie, *Higgs-boson plus Dijets: Higher-Order Matching for High-Energy Predictions*, *JHEP* **08** (2018) 090 [1805.04446].

[2] J.R. Andersen, J.D. Cockburn, M. Heil, A. Maier and J.M. Smillie, *Finite Quark-Mass Effects in Higgs Boson Production with Dijets at Large Energies*, *JHEP* **04** (2019) 127 [1812.08072].

[3] J.R. Andersen, T. Hapola, M. Heil, A. Maier and J. Smillie, *HEJ 2: High Energy Resummation for Hadron Colliders*, *Computer Physics Communications* **245** (2019) 106832 [1902.08430].

[4] J. Andersen, T. Hapola, M. Heil, A. Maier and J. Smillie, *HEJ version 2*, 2019. https://phab.hepforge.org/source/hej/.

[5] J. Andersen, T. Hapola, M. Heil, A. Maier and J. Smillie, "HEJ: High Energy Jets." `https://hej.web.cern.ch/HEJ/`, 2017-2019.

[6] 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.* **B716** (2012) 1 [1207.7214].

[7] CMS collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys.Lett.* **B716** (2012) 30 [1207.7235].

[8] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321.

[9] P.W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508.

[10] G. Guralnik, C. Hagen and T. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585.

[11] OPAL collaboration, *Search for the standard model Higgs boson in e+ e- collisions at s**(1/2) approximately = 192-GeV - 209-GeV*, *Phys. Lett. B* **499** (2001) 38 [hep-ex/0101014].

[12] L3 collaboration, *Standard model Higgs boson with the L3 experiment at LEP*, *Phys. Lett. B* **517** (2001) 319 [hep-ex/0107054].

[13] ALEPH collaboration, *Final results of the searches for neutral Higgs bosons in e+ e- collisions at s**(1/2) up to 209-GeV*, *Phys. Lett. B* **526** (2002) 191 [hep-ex/0201014].

[14] DELPHI collaboration, *Final results from DELPHI on the searches for SM and MSSM neutral Higgs bosons*, *Eur. Phys. J. C* **32** (2004) 145 [hep-ex/0303013].

[15] S. Dittmaier and M. Schumacher, *The Higgs Boson in the Standard Model - From LEP to LHC: Expectations, Searches, and Discovery of a Candidate*, *Prog. Part. Nucl. Phys.* **70** (2013) 1 [1211.4828].

[16] ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, Tevatron Electroweak Working Group, SLD Electroweak, Heavy Flavour Groups collaboration, *Precision Electroweak Measurements and Constraints on the Standard Model*, 1012.2367.

[17] Particle Data Group collaboration, *Review of particle physics*, *J. Phys. G* **37** (2010) 075021.

[18] "The worldwide lhc computing grid." https://home.cern/science/computing/worldwide-lhc-computing-grid.

[19] 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.

[20] V. Gribov and L. Lipatov, *Deep inelastic e p scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438.

[21] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys. B* **126** (1977) 298.

[22] V.S. Fadin, E.A. Kuraev and L.N. Lipatov, *On the Pomeranchuk singularity in asymptotically free theories*, *Phys. Lett.* **B60** (1975) 50.

[23] E.A. Kuraev, L.N. Lipatov and V.S. Fadin, *Multi - Reggeon Processes in the Yang-Mills Theory*, *Sov. Phys. JETP* **44** (1976) 443.

[24] E.A. Kuraev, L.N. Lipatov and V.S. Fadin, *The Pomeranchuk singularity in nonabelian gauge theories*, *Sov. Phys. JETP* **45** (1977) 199.

[25] I.I. Balitsky and L.N. Lipatov, *The Pomeranchuk singularity in quantum chromodynamics*, *Sov. J. Nucl. Phys.* **28** (1978) 822.

[26] A. Buckley et al., *General-purpose event generators for LHC physics*, *Phys. Rept.* **504** (2011) 145 [1101.2599].

[27] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J.* **C76** (2016) 196 [1512.01178].

[28] 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].

[29] Sherpa collaboration, *Event Generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) 034 [1905.09127].

[30] J.R. Andersen and C.D. White, *A New Framework for Multijet Predictions and its application to Higgs Boson production at the LHC*, *Phys. Rev.* **D78** (2008) 051501 [0802.2858].

[31] J.R. Andersen, V. Del Duca and C.D. White, *Higgs Boson Production in Association with Multiple Hard Jets*, *JHEP* **02** (2009) 015 [0808.3696].

[32] J.R. Andersen and J.M. Smillie, *Constructing All-Order Corrections to Multi-Jet Rates*, *JHEP* **01** (2010) 039 [0908.2786].

[33] J.R. Andersen and J.M. Smillie, *The Factorisation of the t-channel Pole in Quark-Gluon Scattering*, *Phys. Rev.* **D81** (2010) 114021 [0910.5113].

[34] J.R. Andersen and J.M. Smillie, *Multiple Jets at the LHC with High Energy Jets*, *JHEP* **06** (2011) 010 [1101.5394].

[35] L.J. Dixon, *A brief introduction to modern amplitude methods*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Particle Physics: The Higgs Boson and Beyond*, pp. 31–67, 2014, 10.5170/CERN-2014-008.31 [1310.5353].

[36] B. Pontecorvo, *Neutrino Experiments and the Problem of Conservation of Leptonic Charge*, *Sov. Phys. JETP* **26** (1968) 984.

[37] Z. Maki, M. Nakagawa and S. Sakata, *Remarks on the unified model of elementary particles*, *Prog. Theor. Phys.* **28** (1962) 870.

[38] G. Apollinari, I. Béjar Alonso, O. Brüning, P. Fessia, M. Lamont, L. Rossi et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1*, .

[39] M.D. Schwartz, *Quantum Field Theory and the Standard Model*, Cambridge University Press (3, 2014).

[40] J.C. Romao and J.P. Silva, *A resource for signs and Feynman diagrams of the Standard Model*, *Int. J. Mod. Phys. A* **27** (2012) 1230025 [1209.6213].

[41] P. Sikivie, L. Susskind, M.B. Voloshin and V.I. Zakharov, *Isospin Breaking in Technicolor Models*, *Nucl. Phys. B* **173** (1980) 189.

[42] M.E. Peskin and D.V. Schroeder, *An Introduction to quantum field theory*, Addison-Wesley, Reading, USA (1995).

[43] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531.

[44] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652.

[45] S. Catani and M. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, *Nucl. Phys. B* **485** (1997) 291 [hep-ph/9605323].

[46] D.A. Kosower, *Antenna factorization of gauge theory amplitudes*, *Phys. Rev. D* **57** (1998) 5410 [hep-ph/9710213].

[47] S. Frixione, *A General approach to jet cross-sections in QCD*, *Nucl. Phys. B* **507** (1997) 295 [hep-ph/9706545].

[48] C. Berger, Z. Bern, L. 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].

[49] 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].

[50] G. Bevilacqua, M. Czakon, M. Garzelli, A. van Hameren, A. Kardos, C. Papadopoulos et al., *HELAC-NLO*, *Comput. Phys. Commun.* **184** (2013) 986 [1110.1499].

[51] G. Cullen, N. Greiner, G. Heinrich, G. Luisoni, P. Mastrolia, G. Ossola et al., *Automated One-Loop Calculations with GoSam*, *Eur. Phys. J. C* **72** (2012) 1889 [1111.2034].

[52] 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].

[53] 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].

[54] J.M. Campbell, R.K. Ellis and C. Williams, *Hadronic production of a Higgs boson and two jets at next-to-leading order*, *Phys.Rev.* **D81** (2010) 074023 [1001.4495].

[55] 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].

[56] 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 (PhysTeV 2019) Les Houches, France, June 10-28, 2019*, 2020 [2003.01700].

[57] Particle Data Group collaboration, *Review of Particle Physics*, *Phys. Rev. D* **98** (2018) 030001.

[58] P. Baikov, K. Chetyrkin and J. Kühn, *Five-Loop Running of the QCD coupling constant*, *Phys. Rev. Lett.* **118** (2017) 082002 [1606.08659].

[59] F. Herzog, B. Ruijl, T. Ueda, J. Vermaseren and A. Vogt, *The five-loop beta function of Yang-Mills theory with fermions*, *JHEP* **02** (2017) 090 [1701.01404].

[60] T. Luthe, A. Maier, P. Marquard and Y. Schroder, *The five-loop Beta function for a general gauge group and anomalous dimensions beyond Feynman gauge*, *JHEP* **10** (2017) 166 [1709.07718].

[61] 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.* **C75** (2015) 132 [1412.7420].

[62] F. Bloch and A. Nordsieck, *Note on the Radiation Field of the electron*, *Phys. Rev.* **52** (1937) 54.

[63] T. Kinoshita, *Mass singularities of Feynman amplitudes*, *J. Math. Phys.* **3** (1962) 650.

[64] T. Lee and M. Nauenberg, *Degenerate Systems and Mass Singularities*, *Phys. Rev.* **133** (1964) B1549.

[65] R. Frederix, S. Frixione, S. Prestel and P. Torrielli, *On the reduction of negative weights in MC@NLO-type matching procedures*, 2002.12716.

[66] HSF Physics Event Generator WG collaboration, *Challenges in Monte Carlo event generator software for High-Luminosity LHC*, 2004.13687.

[67] R. Ellis, W. Stirling and B. Webber, *QCD and collider physics*, vol. 8, Cambridge University Press (2, 2011).

[68] G.P. Salam, *Towards Jetography*, *Eur. Phys. J. C* **67** (2010) 637 [0906.1833].

[69] M. Cacciari, G.P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J.* **C72** (2012) 1896 [1111.6097].

[70] M. Cacciari, G.P. Salam and G. Soyez, *The anti-$k_t$ jet clustering algorithm*, *JHEP* **04** (2008) 063 [0802.1189].

[71] S. Höche, D. Reichelt and F. Siegert, *Momentum conservation and unitarity in parton showers and NLL resummation*, *JHEP* **01** (2018) 118 [1711.03497].

[72] 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].

[73] G. Bewick, S. Ferrario Ravasio, P. Richardson and M.H. Seymour, *Logarithmic accuracy of angular-ordered parton showers*, *JHEP* **04** (2020) 019 [1904.11866].

[74] J.R. Forshaw, J. Holguin and S. Plätzer, *Building a consistent parton shower*, 2003.06400.

[75] M. Dasgupta, F.A. Dreyer, K. Hamilton, P.F. Monni, G.P. Salam and G. Soyez, *Parton showers beyond leading logarithmic accuracy*, 2002.11114.

[76] S. Frixione and B.R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029 [hep-ph/0204244].

[77] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [hep-ph/0409146].

[78] M. Dasgupta, F. Dreyer, G.P. Salam and G. Soyez, *Small-radius jets to all orders in QCD*, *JHEP* **04** (2015) 039 [1411.5182].

[79] M. Dasgupta, F.A. Dreyer, G.P. Salam and G. Soyez, *Inclusive jet spectrum for small-radius jets*, *JHEP* **06** (2016) 057 [1602.01110].

[80] R. Brower, C.E. DeTar and J. Weis, *Regge Theory for Multiparticle Amplitudes*, *Phys. Rept.* **14** (1974) 257.

[81] J.R. Andersen, T. Binoth, G. Heinrich and J.M. Smillie, *Loop induced interference effects in Higgs Boson plus two jet production at the LHC*, *JHEP* **0802** (2008) 057 [0709.3513].

[82] A. Bredenstein, K. Hagiwara and B. Jäger, *Mixed QCD-electroweak contributions to Higgs-plus-dijet production at the LHC*, *Phys.Rev.* **D77** (2008) 073004 [0801.4231].

[83] L.J. Dixon and Y. Sofianatos, *Analytic one-loop amplitudes for a Higgs boson plus four partons*, *JHEP* **0908** (2009) 058 [0906.0008].

[84] J.M. Campbell, R.K. Ellis and G. Zanderighi, *Next-to-Leading order Higgs + 2 jet production via gluon fusion*, *JHEP* **0610** (2006) 028 [hep-ph/0608194].

[85] J.R. Andersen et al., *Les Houches 2017: Physics at TeV Colliders Standard Model Working Report Group*, in *10th Les Houches Workshop on Physics at TeV Colliders (PhysTeV 2017) Les Houches, France, June 5-23, 2017*, 2018, https://inspirehep.net/record/1663483/files/1803.07977.pdf [1803.07977].

[86] V. Del Duca, W. Kilgore, C. Oleari, C. Schmidt and D. Zeppenfeld, *Gluon fusion contributions to H + 2 jet production*, *Nucl.Phys.* **B616** (2001) 367 [hep-ph/0108030].

[87] G. Klämke and D. Zeppenfeld, *Higgs plus two jet production via gluon fusion as a signal at the CERN LHC*, *JHEP* **0704** (2007) 052 [hep-ph/0703202].

[88] ATLAS collaboration, *Measurement of the cross-section for electroweak production of dijets in association with a Z boson in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, *Phys. Lett.* **B775** (2017) 206 [1709.10264].

[89] ATLAS collaboration, *Measurements of electroweak $Wjj$ production and constraints on anomalous gauge couplings with the ATLAS detector*, *Eur. Phys. J.* **C77** (2017) 474 [1703.04362].

[90] ATLAS collaboration, *Measurements of the w production cross sections in association with jets with the atlas detector*, *Eur. Phys. J.* **C75** (2015) 82 [1409.8639].

[91] D0 collaboration, *Studies of W boson plus jets production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, *Phys. Rev.* **D88** (2013) 092001 [1302.6508].

[92] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007 [0811.4622].

[93] J.R. Andersen, V. Del Duca, F. Maltoni and W.J. Stirling, *W boson production with associated jets at large rapidities*, *JHEP* **05** (2001) 048 [hep-ph/0105146].

[94] ATLAS collaboration, *Measurement of dijet production with a veto on additional central jet activity in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, *JHEP* **09** (2011) 053 [1107.1641].

[95] ATLAS collaboration, *Measurements of jet vetoes and azimuthal decorrelations in dijet events produced in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, *Eur. Phys. J.* **C74** (2014) 3117 [1407.5756].

[96] J.R. Forshaw and D. Ross, *Quantum chromodynamics and the pomeron*, vol. 9, Cambridge University Press (1, 2011).

[97] V.S. Fadin, R. Fiore, M.G. Kozlov and A.V. Reznichenko, *Proof of the multi-Regge form of QCD amplitudes with gluon exchanges in the NLA*, *Phys. Lett.* **B639** (2006) 74 [hep-ph/0602006].

[98] J.R. Andersen, T. Hapola, A. Maier and J.M. Smillie, *Higgs Boson Plus Dijets: Higher Order Corrections*, *JHEP* **09** (2017) 065 [1706.01002].

[99] J.R. Andersen, H.M. Brooks and L. Lönnblad, *Merging High Energy with Soft and Collinear Logarithms using HEJ and PYTHIA*, 1712.00178.

[100] V. Del Duca, *Equivalence of the parke-taylor and the fadin-kuraev-lipatov amplitudes in the high-energy limit*, *Phys. Rev.* **D52** (1995) 1527 [hep-ph/9503340].

[101] V.S. Fadin, *BFKL news*, in *LAFEX International School on High-Energy Physics (LISHEP 98) Session A: Particle Physics for High School Teachers - Session B: Advanced School in HEP - Session C: Workshop on Diffractive Physics*, pp. 742–776, 7, 1998 [hep-ph/9807528].

[102] S.J. Parke and T. Taylor, *An Amplitude for n Gluon Scattering*, *Phys. Rev. Lett.* **56** (1986) 2459.

[103] V. Del Duca, *Parke-Taylor amplitudes in the multi - Regge kinematics*, *Phys. Rev.* **D48** (1993) 5133 [hep-ph/9304259].

[104] V. Del Duca, A. Frizzo and F. Maltoni, *Factorization of tree QCD amplitudes in the high-energy limit and in the collinear limit*, *Nucl.Phys.* **B568** (2000) 211 [hep-ph/9909464].

[105] L. Lipatov, *Reggeization of the Vector Meson and the Vacuum Singularity in Nonabelian Gauge Theories*, *Sov. J. Nucl. Phys.* **23** (1976) 338.

[106] V.S. Fadin and L.N. Lipatov, *Next-to-leading corrections to the BFKL equation from the gluon and quark production*, *Nucl. Phys.* **B477** (1996) 767 [hep-ph/9602287].

[107] J.R. Andersen, T. Hapola and J.M. Smillie, *W Plus Multiple Jets at the LHC with High Energy Jets*, *JHEP* **1209** (2012) 047 [1206.6763].

[108] J.R. Andersen, J.J. Medley and J.M. Smillie, *$Z/\gamma^*$ plus multiple hard jets in high energy collisions*, *JHEP* **05** (2016) 136 [1603.05460].

[109] V. Del Duca, W. Kilgore, C. Oleari, C.R. Schmidt and D. Zeppenfeld, *Kinematical limits on Higgs boson production via gluon fusion in association with jets*, *Phys. Rev.* **D67** (2003) 073003 [hep-ph/0301013].

[110] J.R. Andersen, J.A. Black, H.M. Brooks, E.P. Byrne, A. Maier and J.M. Smillie, "*Subleading Corrections to the High Energy Behaviour of W Boson Production in Association with Jets*." Not yet published.

[111] B.L. Ioffe, V.S. Fadin and L.N. Lipatov, *Quantum chromodynamics: Perturbative and nonperturbative aspects*, vol. 30, Cambridge Univ. Press (2010), 10.1017/CBO9780511711817.

[112] A. Bogdan, V. Del Duca, V.S. Fadin and E.N. Glover, *The Quark Regge trajectory at two loops*, *JHEP* **0203** (2002) 032 [hep-ph/0201240].

[113] G. Lepage, *A New Algorithm for Adaptive Multidimensional Integration*, *J. Comput. Phys.* **27** (1978) 192.

[114] Y.L. Dokshitzer, V.A. Khoze and T. Sjöstrand, *Rapidity gaps in Higgs production*, *Phys.Lett.* **B274** (1992) 116.

[115] S. Catani, F. Krauss, R. Kuhn and B.R. Webber, *QCD matrix elements + parton showers*, *JHEP* **11** (2001) 063 [hep-ph/0109231].

[116] L. Lönnblad, *ARIADNE version 4: A Program for simulation of QCD cascades implementing the color dipole model*, *Comput.Phys.Commun.* **71** (1992) 15.

[117] M. Galassi, J. Davies, J. Theiler, B. Gough, G. Jungman, P. Alken et al., *Gnu scientific library reference manual*, 2009. https://www.gnu.org/software/gsl/.

[118] G. Salam. private communication.

[119] E. Boos et al., *Generic user process interface for event generators*, in *Physics at TeV colliders. Proceedings, Euro Summer School, Les Houches, France, May 21-June 1, 2001*, 2001, http://lss.fnal.gov/archive/preprint/fermilab-conf-01-496-t.shtml [hep-ph/0109068].

[120] L. Lonnblad, "C++ parser/writer for les houches event files." `https://gitlab.cern.ch/hepmc/HepMC3/-/blob/master/include/HepMC3/LHEF.h`, 2009–2013.

[121] 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].

[122] Q. Koziol and D. Robinson, *Hdf5*, mar, 2018. 10.11578/dc.20180330.1.

[123] E. Blue Brain Project, "Highfive - hdf5 header-only c++ library." `https://github.com/BlueBrain/HighFive`, 2017–2020.

[124] J. Cruz-Martinez, D. Walker, J. Whitehead and M. Heil, *pyhepgrid: distributed computing made easy*, May, 2019. 10.5281/zenodo.3877494.

[125] M. Ellert, M. Grønager, A. Konstantinov, B. Kónya, J. Lindemann, I. Livenson et al., *Advanced Resource Connector middleware for lightweight computational Grids*, *Future Gener. Comput. Syst.* **23** (2007) 219.

[126] F. Stagni, A. Tsaregorodtsev, ubeda, P. Charpentier, K.D. Ciba, Z. Mathe et al., *Diracgrid/dirac: v6r20p15*, Oct., 2018. 10.5281/zenodo.1451647.

[127] M. Dobbs and J.B. Hansen, *The HepMC C++ Monte Carlo event record for High Energy Physics*, *Comput. Phys. Commun.* **134** (2001) 41.

[128] A. Buckley, P. Ilten, D. Konstantinov, L. Lönnblad, J. Monk, W. Porkorski et al., *The HepMC3 Event Record Library for Monte Carlo Event Generators*, 1912.08005.

[129] A. Buckley, J. Butterworth, L. Lönnblad, D. Grellscheid, H. Hoeth, J. Monk et al., *Rivet user manual*, *Comput. Phys. Commun.* **184** (2013) 2803 [1003.0694].

[130] C. Bierlich et al., *Robust Independent Validation of Experiment and Theory: Rivet version 3*, *SciPost Phys.* **8** (2020) 026 [1912.05451].

[131] G.K. Savvidy and N.G. Ter-Arutunian Savvidy, *ON THE MONTE CARLO SIMULATION OF PHYSICAL SYSTEMS*, *J. Comput. Phys.* **97** (1991) 566.

[132] K.G. Savvidy, *The MIXMAX random number generator*, *Comput. Phys. Commun.* **196** (2015) 161 [1403.5355].

[133] M. Lüscher, *A Portable high quality random number generator for lattice field theory simulations*, *Comput. Phys. Commun.* **79** (1994) 100 [hep-lat/9309020].

[134] L. Lönnblad, *CLHEP: A project for designing a C++ class library for high-energy physics*, *Comput. Phys. Commun.* **84** (1994) 307.

[135] J. Boudreau, G. Cosmo, M. Fischler, L. Garren, B. Jacobsen, N. Katayama et al., "Clhep - a class library for high energy physics." https://gitlab.cern.ch/CLHEP/CLHEP, 2015–2019.

[136] "YAML." https://yaml.org.

[137] J. Beder, "yaml-cpp." https://github.com/jbeder/yaml-cpp/, 2009–2019.

[138] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, *JHEP* **12** (2008) 039 [0808.3674].

[139] J. Kuipers, T. Ueda, J. Vermaseren and J. Vollinga, *FORM version 4.0*, *Comput. Phys. Commun.* **184** (2013) 1453 [1203.6543].

[140] B. Ruijl, T. Ueda and J. Vermaseren, *FORM version 4.2*, 1707.06453.

[141] ATLAS collaboration, *Measurements of fiducial and differential cross sections for Higgs boson production in the diphoton decay channel at* $\sqrt{s} = 8$ *TeV with ATLAS*, *JHEP* **09** (2014) 112 [1407.4222].

[142] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, *Phys. Rev.* **D93** (2016) 033006 [1506.07443].

[143] D.L. Rainwater, R. Szalapski and D. Zeppenfeld, *Probing color singlet exchange in Z + two jet events at the CERN LHC*, *Phys. Rev.* **D54** (1996) 6680 [hep-ph/9605444].

[144] F. Cascioli, P. Maierhöfer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601 [1111.5206].

[145] J. Currie, A. Gehrmann-De Ridder, T. Gehrmann, E.W.N. Glover, A. Huss and J. Pires, *Precise predictions for dijet production at the LHC*, *Phys. Rev. Lett.* **119** (2017) 152001 [1705.10271].

[146] J.M. Campbell et al., *Working Group Report: Quantum Chromodynamics*, in *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*, 2013, http://inspirehep.net/record/1261432/files/arXiv:1310.5189.pdf [1310.5189].

[147] C. Anastasiou, C. Duhr, F. Dulat, E. Furlan, T. Gehrmann, F. Herzog et al., *Higgs boson gluon-fusion production at threshold in $N^3LO$ QCD*, *Phys. Lett.* **B737** (2014) 325 [1403.4616].

[148] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog and B. Mistlberger, *Higgs Boson Gluon-Fusion Production in QCD at Three Loops*, *Phys. Rev. Lett.* **114** (2015) 212001 [1503.06056].

[149] C. Anastasiou, C. Duhr, F. Dulat, E. Furlan, T. Gehrmann, F. Herzog et al., *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, *JHEP* **05** (2016) 058 [1602.00695].

[150] L. Cieri, X. Chen, T. Gehrmann, E.W.N. Glover and A. Huss, *Higgs boson production at the LHC using the $q_T$ subtraction formalism at $N^3LO$ QCD*, 1807.11501.

[151] R.V. Harlander and K.J. Özeren, *Finite top mass effects for hadronic Higgs production at next-to-next-to-leading order*, *JHEP* **11** (2009) 088 [0909.3420].

[152] R.V. Harlander, H. Mantler, S. Marzani and K.J. Özeren, *Higgs production in gluon fusion at next-to-next-to-leading order QCD for finite top mass*, *Eur. Phys. J.* **C66** (2010) 359 [0912.2104].

[153] R.V. Harlander, T. Neumann, K.J. Özeren and M. Wiesemann, *Top-mass effects in differential Higgs production through gluon fusion at order $\alpha_s^4$*, *JHEP* **08** (2012) 139 [1206.0157].

[154] S.P. Jones, M. Kerner and G. Luisoni, *NLO QCD corrections to Higgs boson plus jet production with full top-quark mass dependence*, *Phys. Rev. Lett.* **120** (2018) 162001 [1802.00349].

[155] K. Becker et al., *Precise predictions for boosted Higgs production*, 2005.07762.

[156] J.M. Lindert, K. Melnikov, L. Tancredi and C. Wever, *Top-bottom interference effects in Higgs plus jet production at the LHC*, *Phys. Rev. Lett.* **118** (2017) 252002 [1703.03886].

[157] J.R. Andersen, K. Arnold and D. Zeppenfeld, *Azimuthal Angle Correlations for Higgs Boson plus Multi-Jet Events*, *JHEP* **1006** (2010) 091 [1001.3822].

[158] M. Cacciari, F.A. Dreyer, A. Karlberg, G.P. Salam and G. Zanderighi, *Fully Differential Vector-Boson-Fusion Higgs Production at Next-to-Next-to-Leading Order*, *Phys. Rev. Lett.* **115** (2015) 082002 [1506.02660].

[159] 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.* **B781** (2018) 672 [1802.02445].

[160] F.A. Dreyer and A. Karlberg, *Vector-Boson Fusion Higgs Production at Three Loops in QCD*, *Phys. Rev. Lett.* **117** (2016) 072001 [1606.00840].

[161] V. Del Duca, W. Kilgore, C. Oleari, C. Schmidt and D. Zeppenfeld, *Higgs + 2 jets via gluon fusion*, *Phys.Rev.Lett.* **87** (2001) 122001 [hep-ph/0105129].

[162] L. Budge, J.M. Campbell, G. De Laurentis, R.K. Ellis and S. Seth, *The one-loop amplitudes for Higgs + 4 partons with full mass effects*, *JHEP* **05** (2020) 079 [2002.04018].

[163] G. Cullen, H. van Deurzen, N. Greiner, G. Luisoni, P. Mastrolia, E. Mirabella et al., *Next-to-Leading-Order QCD Corrections to Higgs Boson Production Plus Three Jets in Gluon Fusion*, *Phys. Rev. Lett.* **111** (2013) 131801 [1307.4737].

[164] N. Greiner, S. Höche, G. Luisoni, M. Schönherr and J.-C. Winter, *Full mass dependence in Higgs boson production in association with jets at the LHC and FCC*, *JHEP* **01** (2017) 091 [1608.01195].

[165] S. Carrazza, R.K. Ellis and G. Zanderighi, *QCDLoop: a comprehensive framework for one-loop scalar integrals*, *Comput. Phys. Commun.* **209** (2016) 134 [1605.03181].

[166] A.H. Hoang and I.W. Stewart, *Top Mass Measurements from Jets and the Tevatron Top-Quark Mass*, *Nucl. Phys. Proc. Suppl.* **185** (2008) 220 [0808.0222].

[167] J. Kieseler, K. Lipka and S.-O. Moch, *Calibration of the Top-Quark Monte Carlo Mass*, *Phys. Rev. Lett.* **116** (2016) 162001 [1511.00841].

[168] M. Butenschoen, B. Dehnadi, A.H. Hoang, V. Mateu, M. Preisser and I.W. Stewart, *Top Quark Mass Calibration for Monte Carlo Event Generators*, *Phys. Rev. Lett.* **117** (2016) 232001 [1608.01318].

[169] A.H. Hoang, S. Mantry, A. Pathak and I.W. Stewart, *Extracting a Short Distance Top Mass with Light Grooming*, 1708.02586.

[170] P. Nason, *The Top Mass in Hadronic Collisions*, in *From My Vast Repertoire ...: Guido Altarelli's Legacy*, A. Levy, S. Forte and G. Ridolfi, eds., pp. 123–151 (2019), 10.1142/9789813238053_0008 [1712.02796].

[171] A.H. Hoang, S. Plätzer and D. Samitz, *On the Cutoff Dependence of the Quark Mass Parameter in Angular Ordered Parton Showers*, *JHEP* **10** (2018) 200 [1807.06617].

[172] CMS collaboration, *Measurement of the top quark mass using proton-proton data at* $\sqrt{(s)}$ *= 7 and 8 TeV*, *Phys. Rev.* **D93** (2016) 072004 [1509.04044].

[173] ATLAS collaboration, *Measurement of the top quark mass in the* $t\bar{t} \to$ *dilepton channel from* $\sqrt{s} = 8$ *TeV ATLAS data*, *Phys. Lett.* **B761** (2016) 350 [1606.02179].

[174] CDF AND D0 collaboration, *Combination of CDF and D0 results on the mass of the top quark using up* 9.7 $\text{fb}^{-1}$ *at the Tevatron*, 1608.01881.

[175] U. Langenfeld, S. Moch and P. Uwer, *Measuring the running top-quark mass*, *Phys. Rev.* **D80** (2009) 054009 [0906.5273].

[176] D0 collaboration, *Determination of the pole and* $\overline{MS}$ *masses of the top quark from the* $t\bar{t}$ *cross section*, *Phys. Lett.* **B703** (2011) 422 [1104.2887].

[177] P. Marquard, A.V. Smirnov, V.A. Smirnov and M. Steinhauser, *Quark Mass Relations to Four-Loop Order in Perturbative QCD*, *Phys. Rev. Lett.* **114** (2015) 142002 [1502.01030].

[178] T. Plehn, D.L. Rainwater and D. Zeppenfeld, *Determining the structure of Higgs couplings at the LHC*, *Phys. Rev. Lett.* **88** (2002) 051801 [hep-ph/0105325].

[179] L.J. Dixon, *Calculating scattering amplitudes efficiently*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 95): QCD and Beyond*, pp. 539–584, 1, 1996 [hep-ph/9601359].