Department of Physics and Astronomy

University of Heidelberg

Diploma thesis

in Physics

submitted by

Thomas Lübbert

born in **Münster**

October 2010

Resummation of event shapes
at hadron colliders

This diploma thesis has been carried out by

Thomas Lübbert

at the

Institute for Theoretical Physics

under the supervision of

Prof. Tilman Plehn

Abstract

We discuss two complementary approaches to account for the leading logarithms of QCD calculations to all orders in perturbation theory: The Monte Carlo parton shower and the method of resummation. These approaches are implemented in the programs SHERPA [1] and CAESAR [2]. We show that the integrated form of the splitting probabilities used in SHERPA have next-to-leading logarithmic (NLL) accuracy in the limit of massless partons. We present results obtained with SHERPA for distributions of a range of dijet event shapes at the Tevatron collider. These results will be compared to predictions of NLO+NLL accuracy obtained by CAESAR in [3]. We find agreement between the two complementary approaches within theoretical uncertainties.

Contents

Contents

1 Introduction

Particle physics aims to describe the properties of matter and its interactions down to very small distances. The currently best established theory of elementary particle physics, the Standard Model, predicts, for reasons of self-consistency, the existence of a yet unobserved particle, the Higgs boson. While direct searches give a lower limit on its mass of order 100 GeV, electroweak precision data from the LEP experiments limit its mass from above by about 200 GeV. Hence, if the Standard Model is correct, the Higgs boson should be observable in this energy range. Moreover, many models for physics beyond the Standard Model predict additional heavy particles, which might be discovered at high energies. Tests for the existence of the these particles and the Higgs boson, to falsify or support the corresponding theories, require a sufficiently high center of mass energy in scattering experiments. The only accelerators that provide such high energies are the hadron-hadron colliders "Tevatron" at Fermilab and "LHC" at CERN. Since they collide hadrons, whose constituents, the partons, are color charged, most of the scattering events that occur involve only color charged particles which are described by the theory of Quantum Chromodynamics (QCD).

For new physics searches, however, these have to be considered as a background and need to be reduced. Thus, to describe the physics of hadron colliders and eventually extract signals of new physics, one needs to understand QCD very well and describe its observable implications with high accuracy. For this goal, fixed order calculations turn out to be insufficient, but methods are required that resum the dominant effects of higher order corrections to all orders. Two of these methods will be the main subjects of this work. Besides discussing their theoretical foundations, their predictions for the distribution of dijet event shape observables will be studied in detail. Event shape variables measure the geometrical properties of the momentum flow in events and, notably, its deviation from pure lowest order predictions. Hence, they provide a laboratory to test perturbative QCD, both its fixed order and its resummation approaches. This is our main motivation to look at them. Moreover, they might serve as tools for new physics searches, as the energy flow of most new physics events deviates considerably from pure QCD events. While event shapes for other collider types are know to high accuracy, at hadron colliders still a better understanding is needed. A contribution in this direction has been given in [3].

The structure of this work will be as follows. In chapter 2, we will introduce the relevant theory of QCD in form of its Lagrangian and show, how it is used to obtain experimentally testable predictions. Along the way, we will encounter difficulties related to strongly enhanced or even divergent contributions. We will introduce the basic concepts that allow to control these difficulties. Related to this, we will look at general properties of QCD matrix elements and the multi particle phase space in chapter 3. We will find that the

emission of soft and collinear particles is strongly enhanced and that its leading effects can be factorized from a hard core-process for hard scattering events.

Using this factorization property, we will discuss methods to account for the dominant effects of this radiation to all orders in perturbation theory in chapters 3.4, 4 and 6. For the methods of Monte Carlo parton showers and resummation, we consider explicit examples of programs, which are SHERPA [1] and CAESAR [2], respectively. While those topics are known in the literature, we extend the understanding of the accuracy of SHERPA in this work. In section 3.3, we show that the shower ordering parameter used in SHERPA respects next-to-leading logarithmic (NLL) accuracy as defined later. In chapter 5, we proof the same accuracy for the integrated form of the splitting probabilities used in SHERPA. During reviewing the discussion of resummation, we find that SHERPA itself can provide NLL accuracy only in the limit of a large number of colors ($N_c \to \infty$). In chapter 8, we will compare predictions from SHERPA and CAESAR for the event shape observables, which we already mentioned above and which will be defined in chapter 7. In chapter 9, we will conclude.

2 General aspects

In the first section of this chapter, we will introduce the theory of Quantum Chromodynamics (QCD), which describes the interaction of the constituents of hadrons and which is the theory discussed throughout this work. In section 2.2, we will outline, how a quantum field theory as QCD can be used to predict experimentally measurable observables by perturbative calculations, and what kinds of problems are encountered on the way.

2.1 Quantum Chromodynamics

The current theory describing elementary particle physics is the Standard Model (SM). It contains fermions, i.e. quarks and leptons, as well as bosons, i.e. gauge bosons and Higgs bosons. The gauge bosons transmit interactions between the appropriately charged fermions. They have to be introduced, to make the Lagrangian invariant under the $SU(3)_C \times SU(2)_L \times U(1)_{em}$ gauge group, which is a local symmetry of the theory. The strength of various gauge interactions is governed by the corresponding coupling constant α.

For particles charged under $SU(3)_C$, i.e. they carry color charge, the "strong interaction" is the dominant one. It is transmitted by the gluons, the gauge bosons of $SU(3)_C$, and it is described by the theory of QCD. In this work, we will discuss jet physics at hadron colliders. Since hadrons consist of color charged constituents, which are (anti)quarks ($^{(}\bar{q}^{)}$) and gluons (g), we will concentrate on QCD in the following.

$SU(3)$ Symmetry

Quarks are fermions, i.e. spinors with respect to Lorentz symmetry, which, of course, is a symmetry of the QCD Lagrangian. Quarks are in the fundamental ("3") representation of $SU(3)$, i.e. in color space they are three-vectors labeled by the color index $a = 1, 2, 3$ which transform as

$$q_a(x) \rightarrow q'_a(x) = U_{ab}(x)q_b(x)\,, \tag{2.1}$$

where $U_{ab}(x)$ is in the fundamental representation of $SU(3)$, i.e. a unitary 3×3 matrix with determinant 1. In terms of the $8 = 3^2 - 1$ generators t^A, $A = 1, ..., 8$ of $SU(3)$, we can write

$$U(x) = e^{-\alpha^A(x)t^A}\,, \tag{2.2}$$

where, as in eq. (2.1), a sum over repeated indices is understood. The generators t^A are hermitian and traceless matrices as implied by the corresponding requirements on U. They fulfill the relations

$$[t_r^A, t_r^B] = f^{ABC} t_r^C \,, \tag{2.3}$$

$$\mathrm{tr}(t_r^A t_r^B) = T_r \delta^{AB} \quad \text{with } T_F = \frac{1}{2}, \ T_A = 3 \,, \tag{2.4}$$

$$\sum_A t_r^A t_r^A = C_r \, 1_{d(r)} \,. \tag{2.5}$$

These relations hold in any representation r of the group. The constant C_r in eq. (2.5) is called Casimir operator of the representation r. The constant T_r is related to it by

$$T_r = \frac{d(r)}{d(A)} C_r \,, \tag{2.6}$$

where A labels the adjoint representation. The latter is $d(A) = 8$ dimensional, and fulfills the special relations $(t^B)^{CD} = -i f^{BCD}$ and $T_A = C_A$. In eq. (2.4), we fixed the normalization of t_r^A and f^{ABC} by the choice $T_A = 3$.

In eq. (2.3), f^{ABC} are the structure constants of $SU(3)$, which are totally antisymmetric and independent of the representation. The non-vanishing structure constants that are not related by symmetries are

$$f^{147} = f^{156} = f^{246} = f^{257} = f^{345} = f^{367} = \frac{1}{2},$$

$$f^{123} = 1, \ f^{458} = f^{678} = \frac{\sqrt{3}}{2} \,. \tag{2.7}$$

Lagrangian of QCD

To get the QCD Lagrangian, we start with the Dirac Lagrangian

$$\mathcal{L}_D = \bar{\psi}(x) \left(i \, \slashed{\partial} - m \right) \psi(x) \,, \tag{2.8}$$

where ψ represents a quark and $\slashed{\partial} = \partial_\mu \gamma^\mu$ with the Lorentz gamma matrices γ^μ. Performing a local $SU(3)$ transformation as in eq. (2.1), this transforms as

$$\bar{\psi} \left(i \, \slashed{\partial} - m \right) \psi \rightarrow \mathcal{L}_D + \bar{\psi}_a \left[i U_{ab}^\dagger(x) \partial_\mu U_{bc}(x) \right] \gamma^\mu \psi_c \,. \tag{2.9}$$

Thus, eq. (2.8) does not obey local $SU(3)$ symmetry. To restore the symmetry, we need to add an additional term, which transforms as minus the second term of eq. (2.9). Since again it has to contain ψ and $\bar{\psi}$, we only need to look at the expression in the squared bracket in eq. (2.9), which carries a vector index μ, two color indices a,c and depends on x. Therefore, we introduce the fields $A_{\mu\,ab}(x) = A_\mu^A(x) t_{ab}^A$, which interact with the fermions in the following way

$$\mathcal{L}_I = g_S \bar{\psi}_a A_{\mu\,ab} \gamma^\mu \psi_b \,. \tag{2.10}$$

The fields $A_\mu^A(x)$ are called gluons. In order to guarantee gauge invariance of $\mathcal{L}_D + \mathcal{L}_I$, $A_{\mu\,ab}(x)$ has to transform in the appropriate way. Under Lorentz symmetry, it transforms as a vector. In color space, where it is a matrix, it has to transform as

$$A_{\mu\,ab}(x) \rightarrow U_{ac}(x)A_{\mu\,cd}(x)U_{db}^\dagger(x) - \frac{i}{g_S}\left(U_{ac}^\dagger(x)\partial_\mu U_{cb}(x)\right). \tag{2.11}$$

With this choice, $\mathcal{L}_D + \mathcal{L}_I$ is obviously gauge invariant. Introducing the covariant derivative

$$D_{\mu\,ab} = \partial_\mu \delta_{ab} - ig_S A_{\mu\,ab}(x), \tag{2.12}$$

we can write the Lagrangian as

$$\mathcal{L}_D + \mathcal{L}_I = \bar{\psi}\left(i\,\slashed{D} - m\right)\psi, \tag{2.13}$$

where we suppressed the color indices as we did for the spinor indices from the very beginning. Only one other invariant and renormalizable term, which is build from ψ and A_μ, is relevant. It is

$$\mathcal{L}_G = -\frac{1}{2}\text{tr}\left(F_{\mu\nu}F^{\mu\nu}\right) = -\frac{1}{4}F_{\mu\nu}^A F^{A\mu\nu}, \tag{2.14}$$

where $F_{\mu\nu}$ is defined as

$$F_{\mu\nu} = \frac{i}{g_S}[D_\mu, D_\nu]. \tag{2.15}$$

Using the relation eq. (2.3), this can be expressed in terms of the gluon fields and the $SU(3)$ generators as

$$F_{\mu\nu,ab} = F_{\mu\nu}^A t_{ab}^A \quad \text{with} \tag{2.16}$$
$$F_{\mu\nu}^A = \partial_\mu A_\nu^A - \partial_\nu A_\mu^A + g_S f^{ABC} A_\mu^B A_\nu^C. \tag{2.17}$$

Our discussion so far was only for a single kind of quark. In nature have been observed six flavors of quarks [4] which are down, up, strange, charm, bottom and top. For each of them there is a term as eq. (2.13) with ψ replaced by ψ_{flavor}. Adding these terms and $\mathcal{L}_G$, we arrive at the classical QCD Lagrangian

$$\boxed{\mathcal{L}_{QCD}^{cl} = \sum_f \bar{\psi}_f(i\,\slashed{D} - m_f)\psi_f - \frac{1}{4}F_{\mu\nu}^A F^{A\mu\nu}.} \tag{2.18}$$

Including all indices and sums explicitly, this is

$$\mathcal{L}_{QCD}^{cl}(x) = \underbrace{\sum_{f=1}^{6}}_{\text{flavor}}\underbrace{\sum_{\mu=0}^{3}}_{\text{lorentz}}\underbrace{\sum_{\alpha\beta=1}^{4}}_{\text{spinor}}\underbrace{\sum_{a,b=1}^{3}}_{\text{color}}\bar{\psi}_{f,a,\alpha}(x)\left(iD_{\mu,ab}(x)\gamma_{\alpha\beta}^\mu - m_f\delta_{ab}\delta_{\alpha\beta}\right)\psi_{f,b,\beta}(x)+ \tag{2.19}$$

$$-\underbrace{\sum_{\mu,\nu=0}^{3}}_{\text{lorentz}}\underbrace{\sum_{A=1}^{8}}_{\text{gluon}}\tfrac{1}{4}F_{\mu\nu}^A(x)F^{A\mu\nu}(x).$$

Since this is difficult to read, we will suppress spinor and color indices in the following.

The usual way to quantize a theory is to impose (anti)commutator relations for the fields or use the path integral formalism (see e.g. [5], [6]). In case of the classical QCD Lagrangian, we will run into trouble. In the path integral formalism, for example, the generating functional will diverge. The divergence comes from the integral over the gauge field, which is performed not only over really distinct A_μ^A, but also over the various A_μ^As that are related by gauge symmetry. To cure this problem following Faddeev and Popov [7], we introduce a gauge fixing term in the Lagrangian

$$\mathcal{L}_{fix} = -\frac{1}{2\lambda}\left(\partial^\mu\partial^\nu A_\mu^A(x)A_\nu^A(x)\right) - \bar{c}^A(x)\partial^\mu D_\mu^{AC}c^C(x)\,, \tag{2.20}$$

where λ is the gauge fixing parameter and we introduced the new (non-physical) fields c^A called Fadeev-Popov-ghosts. These are complex scalar fields carrying color. They appear only in loops. The covariant derivative in the adjoint representation is

$$D_\mu^{AB} = \partial_\mu - g_S f^{ABC} A_\mu^C\,. \tag{2.21}$$

The Lagrangian

$$\mathcal{L}_{QCD}^{FP} = \mathcal{L}_{QCD}^{cl} + \mathcal{L}_{fix} \tag{2.22}$$

can be quantized leading to the Feynman rules given in Figure 2.1. The first three are propagators and the following four are vertices. Using the Feynman rules, we can evaluate Feynman diagrams in terms of which matrix elements needed in calculations of cross sections can be expressed perturbatively.

2.2 Perturbation theory

In this section, we will recall, how a quantum field theory as QCD can be used to predict experimentally measurable observables by perturbative calculations. In these calculations, divergences of various kinds arise. We will outline the types of these divergences and the concepts, which are needed to control them.

Cross section

Given any QCD scattering process $a \to b$ with some initial state a and final state b, the probability for this scattering to occur is given by the cross section

$$\sigma_{a\to b} = \frac{1}{f_a} \cdot \left(\prod_{f\in b}\int\frac{\mathrm{d}^3 p_f}{(2\pi)^3 2E_f}\right) \cdot |\mathcal{M}_{a\to b}|^2 \cdot \delta^4\big(\textstyle\sum_{i\in a}p_i - \sum_{f\in b}p_f\big)\,. \tag{2.23}$$

The first factor is the initial state flux. For two initial particles it is given by

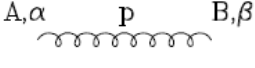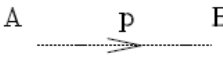$$\frac{1}{f_a} = \frac{1}{2E_1 2E_2|v_1 - v_2|}\,. \tag{2.24}$$
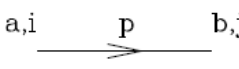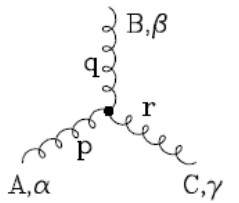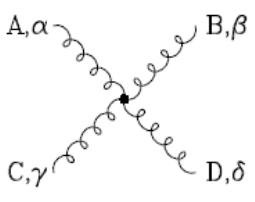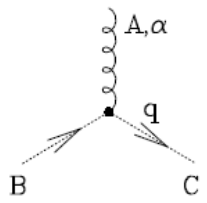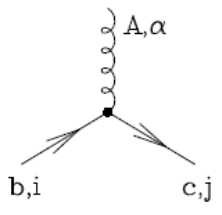
$$\delta^{AB} \left[-g^{\alpha\beta}+(1-\lambda)\frac{p^{\alpha}p^{\beta}}{p^2+i\epsilon}\right]\frac{i}{p^2+i\epsilon}$$

$$\delta^{AB}\frac{i}{(p^2+i\epsilon)}$$

$$\delta^{ab}\frac{i}{(\not{p}-m+i\epsilon)_{ji}}$$

$$-g\ f^{ABC}[(p-q)^{\gamma}g^{\alpha\beta}+(q-r)^{\alpha}g^{\beta\gamma}+(r-p)^{\beta}g^{\gamma\alpha}]$$

(all momenta incoming, $p+q+r = 0$)

$$-ig^2\ f^{XAC}f^{XBD}\ [g^{\alpha\beta}g^{\gamma\delta}-g^{\alpha\delta}g^{\beta\gamma}]$$
$$-ig^2\ f^{XAD}f^{XBC}\ [g^{\alpha\beta}g^{\gamma\delta}-g^{\alpha\gamma}g^{\beta\delta}]$$
$$-ig^2\ f^{XAB}f^{XCD}\ [g^{\alpha\gamma}g^{\beta\delta}-g^{\alpha\delta}g^{\beta\gamma}]$$

$$g\ f^{ABC}q^{\alpha}$$

$$-ig\ (t^A)_{cb}\ (\gamma^{\alpha})_{ji}$$

Figure 2.1: Feynman rules for QCD in covariant gauge. Two further rules are: conserve four-momentum at each vertex and integrate over undetermined four-momenta. In the figure quarks are represented by solid, gluons by curly and ghosts by dotted lines. p, q and r label momentum, A, B, C and a, b label color in the adjoint and the fundamental representation respectively, $\alpha, \beta, \gamma, \delta$ are Lorentz indices labeling the polarization of the gluon. Figure extracted from [8].

The second factor is the phase space of the final state. Removing integrals in eq. (2.23), but keeping the corresponding integration measures, we get a differential cross section. The third factor is the absolute value squared of the invariant matrix element $\mathcal{M}_{a \to b}$. The fourth factor corresponds to total four-momentum conservation.

The invariant matrix element $\mathcal{M}$ can perturbatively be expanded in powers of the strong coupling g_s, given the latter is sufficiently small, i.e. $\frac{g_S^2}{4\pi} < 1$. Then $\mathcal{M}$ can be expressed as the sum of all (connected, amputated and distinct) Feynman diagrams, which can be drawn for the given initial lines a and external lines b. Each diagram is build from the pieces shown in Figure 2.1 and is a shorthand notation for the corresponding (algebraic) term. To connect several lines, we have to include the appropriate vertices from Figure 2.1. Each vertex contributes a factor of the perturbation parameter g_S (the four-gluon-vertex a factor of g_S^2). Hence, each Feynman diagram is suppressed by some power of g_S. A perturbative expansion of $\mathcal{M}$ in g_S means to represent $\mathcal{M}$ as

$$\mathcal{M}_{a \to b} = g_S^n \sum_{k=0}^{\infty} g_S^{2k} \tilde{\mathcal{M}}_{a \to b}^{(n+2k)} \tag{2.25}$$

and abort the sum at some order k of the perturbation. All powers of g_S have been extracted from $\tilde{\mathcal{M}}$. In the truncated series, only a finite number of diagrams will appear, since n+k (the maximal number of vertices) is finite. These diagrams we can straightforwardly calculate by using the Feynman rules. The effort to do this will drastically increase with each order.

Naively, one may expect that the next order in g_S is only suppressed by one not by two powers of g_S. However, the number of external legs is fixed. Increasing the power of g_S by one means to include one additional three-vertex, but than one vertex would have an end not connected to any line. Thus, we have to include two additional three-vertices or one four-vertex. Both correspond to raising the power of g_S by two. In addition to g_S^2 there usually appear factors of $\frac{1}{4\pi}$. Therefore, one chooses

$$\alpha_S = \frac{g_S^2}{4\pi} \tag{2.26}$$

as perturbation parameter, i.e.

$$\mathcal{M}_{a \to b} = \alpha_S^{n/2} \sum_{k=0}^{\infty} \alpha_S^k \mathcal{M}_{a \to b}^{(n/2+k)} \, . \tag{2.27}$$
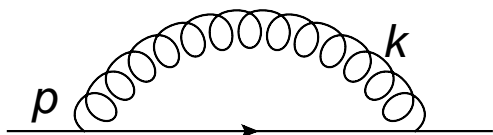
The term with the lowest power of α_S is the leading order (LO) expression; the summand with an additional factor α_S is the next to leading order (NLO) correction and the term suppressed by a factor of α_S^k relative to the LO expression is the N^kLO contribution.

Divergences

In the LO diagrams all internal propagators have usually a fixed momentum given by four-momentum conservation at each vertex. Since the corresponding diagrams look like

branches of trees, they are called tree level diagrams. The higher order diagrams in general include loops, i.e. some internal lines are linked in a way, that they form a closed loop. Adding a four-momentum k to the four-momenta of all loop propagators still preserves momentum conservation at each vertex. Therefore, we have to integrate over all values of k. Thus, loops create expressions like

$$\sim \int \frac{\mathrm{d}^4 k}{(2\pi)^4} \gamma^\mu \frac{\not{p} + \not{k} + m}{(p+k)^2 - m^2 + i\epsilon} \gamma_\mu \frac{1}{k^2 + i\epsilon} .$$

When we try to perform such integrals, problems will arise. First of all the propagators have poles, when the propagating particle is on mass-shell (i.e. $p_i^2 = m_i^2$). Thanks to the $+i\epsilon$ prescription, we avoid the actual pole, but merely encourage the virtual particle to be on-shell. However, if k vanishes or is collinear to p (in the example above), several denominators can vanish (up to the $i\epsilon$) at the same time and we may find divergences. They are called virtual infrared (IR) and collinear (C) divergences. "Virtual", as they arise from propagators of virtual particles.

The infrared and collinear (IRC) sector will be the main focus of this work. We will mostly discuss methods to account for the emission of real, i.e. external, particles. For them, too, IRC divergences arise. The crucial observation, allowing us to remove the singularities later, is that the real and virtual IRC singularities cancel for appropriately inclusive observables [9, 10, 11]. "Inclusive" means that the emission of additional external particles, which are very soft or collinear to another particle, should not change the value of the observable much.

There is still another problem with integrals like the one above. That is when k'[1] gets large. Then the other momentum and mass scales are negligible and we naively expect the integral above to be

$$\lim_{\Lambda \to \infty} \int_0^\Lambda \mathrm{d}k' k'^3 \frac{k'}{k'^4} = \lim_{\Lambda \to \infty} k' \Big|_0^\Lambda , \tag{2.28}$$

which diverges strongly. [2] Divergences like these are called ultraviolet (UV) divergences. A first step to deal with them is to regulate and quantify them. A method to achieve this is dimensional regularization (see [5]). There the integral over the loop momenta is not done in four, but in $d = 4 - 2\epsilon$ space-time dimensions. One then gets a finite piece, a term proportional to $1/\epsilon$ and terms that vanish for $\epsilon \to 0$.

One can absorb these $1/\epsilon$ terms by the method of renormalization. To do so, we reinterpret the quantities g_s, m, ψ, A_μ in our Lagrangian as "bare" quantities, which differ from

---

[1]We use a Wick-rotation to get the Euclidean version of k.

[2]Actually the divergence of our diagram is only like $\log(\Lambda)$, since the k' in the numerator vanishes in the integral due to symmetry.

the corresponding physical quantities by a "Z factor"

$$g_{s,0} = Z_{g_s} \cdot g_{s,\,\text{phys}} \,. \tag{2.29}$$

Writing then $Z_{g_s} = 1 + (Z_{g_s} - 1)$ etc. we get the old Lagrangian with the bare quantities replaced by the physical ones plus counterterms containing the Z factors. The counterterms introduce some additional Feynman rules, whose contribution we want to cancel the divergences. This determines the counterterms and from that the Z factors. The nontrivial fact that this is possible with these fixed number of Z factors and therefore a finite number of counterterms at all orders of perturbation theory, means that QCD is "renormalizable" and follows from the fact that α_S is dimensionless and the form of our Lagrangian.

Running coupling constant

In the regularization step, we are forced by dimensional reasons to introduce a new mass scale - the renormalization scale μ_r - in our Lagrangian. It appears in various terms and hence also in the expressions for the formerly divergent integrals, that are regulated now. Therefore, the Z factors will depend on that scale, too. Since the bare parameters cannot depend on this arbitrary scale, our physical parameters have to depend on it. This dependence is described by differential equations called renormalization group equations.

The one, most important to us, is the dependence of the renormalized coupling "constant" on the renormalization scale, which is described by

$$\beta(\alpha_S(\mu_r^2)) = \mu_r^2 \frac{\partial \alpha_S(\mu_r^2)}{\partial \mu_r^2} = \frac{\partial \alpha_S(\mu_r^2)}{\partial \log \mu_r^2} \,. \tag{2.30}$$

The perturbative expansion of β is

$$\beta(\alpha_S) = -\alpha_S \sum_{n=0}^{\infty} \beta_n \left(\frac{\alpha_S}{4\pi} \right)^{n+1} \,. \tag{2.31}$$

Here in the $\overline{MS}$ scheme and with n_f the number of active quark flavors we have up to third perturbative order [12]

$$\beta_0 = 11 - \frac{2}{3} n_f, \quad \beta_1 = 102 - \frac{38}{3} n_f \,, \tag{2.32}$$

$$\beta_2 = \frac{2857}{2} - \frac{5033}{18} n_f + \frac{325}{54} n_f^2 \,. \tag{2.33}$$

Then the running coupling constant to second perturbative order is given by

$$\alpha_S(Q^2) = \frac{\alpha_S(\mu^2)}{1 + \frac{\beta_0}{4\pi} \alpha_S(\mu^2) \log \frac{Q^2}{\mu^2}} \left[1 - \frac{\beta_1}{4\pi\beta_0} \frac{\alpha_S(\mu^2) \log(1 + \frac{\beta_0}{4\pi} \alpha_S(\mu^2) \log \frac{Q^2}{\mu^2})}{1 + \frac{\beta_0}{4\pi} \alpha_S(\mu^2) \log \frac{Q^2}{\mu^2}} \right]$$
$$+ \mathcal{O}\left(\alpha_S^{k+3} \log^k \frac{Q^2}{\mu^2} \right) \,. \tag{2.34}$$
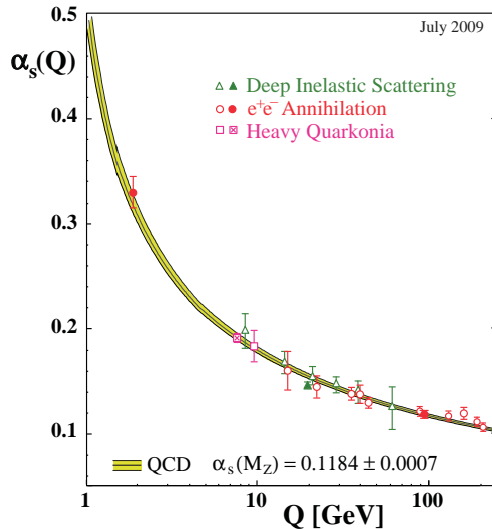
Figure 2.2: The running of the strong coupling constant. The theory prediction (yellow band) is compared to measurements from various experiments (markers). Figure extracted from [4] .

That is, choosing Q^2 instead of μ^2 as the renormalization scale, the renormalized coupling changes in a well defined manner. The running of α_S is also shown in Figure 2.2, where the theory prediction is compared to measurements from various experiments. Although the choice of μ_r is arbitrary and in fact for physical observables calculated to all orders its dependence cancels, some choices for μ_r are more convenient than others, since corrections from higher orders are smaller for them. A good choice of μ_r is the typical momentum scale Q of the process. In hadron-hadron collision this can be the transverse momentum of the final jets; in deep inelastic scattering (DIS) the typical scale is the transferred momentum. In an expression for an observable in terms of $\alpha_S(\mu_r)$, logarithms of Q^2/μ_r^2 will appear with each power of α_S. If we choose μ_r^2 very different from Q^2, these logarithms are large and enhance higher order corrections. To keep higher order corrections small, it is, therefore, best to account for them by choosing a proper scale for α_S.

We observe, α_S in eq. (2.34) decreases, when Q^2 increases. This means, for processes at a sufficiently large scale Q^2 we will have $\alpha_S(Q^2) \ll 1$ and perturbation theory is well justified there. However, for $Q < \mu$ the logarithm in eq. (2.34) will become negative. At some point $Q = \Lambda_{QCD}$ the denominator will vanish giving $\alpha_S(\Lambda_{QCD}^2) \to \infty$. Thus, perturbation theory is not justified there. Hence, at low Q non-perturbative effects will be important. One such effect is confinement: We do not observe free partons, i.e. quarks or gluons, but in free particles several of them are combined to form a color singlet called hadron.[3]

---

[3]Since we do not observe quarks and gluons directly, one may be worried there is no evidence for them in nature. However, scattering experiments involving hadrons have clearly shown, that they are not elementary [13, 14], but are build from constituents. Moreover, the spectrum of the hadrons hints to the symmetry of QCD [4]. Many other experiments have been performed all agreeing with theoretical predictions .

In a simplified picture, this can be seen as follows. When we try to separate a quark and an antiquark of the corresponding anticolor, with their increasing distance r, the relevant $\alpha_S(Q^2)$ would increase, too, since roughly $Q^2 \sim 1/r^2$ - recall De Broglie's relation $\lambda = h/p$. The potential energy of the color dipole then grows $\sim r$. At some point, it is large enough to create a quark-antiquark pair from the vacuum. Thus, instead of a separated quark and antiquark, we now have two color neutral quark-antiquark pairs, which can form hadrons.

Factorization

To summarize, we discussed how to calculate scattering amplitudes between partons in perturbation theory. However, in an experiment we cannot observe partons directly or prepare an isolated parton in the initial state. Instead, hadrons will appear there, whose formation out of partons cannot be calculated perturbatively, since this happens at low momentum scales, where α_S is large.

Is there a way to use perturbative QCD to predict observables in scattering experiments? There are two essential observations helping here, which are factorization [15] and asymptotic freedom [16,17]. At small distances (large Q^2), α_S is small and the partons inside the hadrons are, therefore, quasi free. This is called asymptotic freedom. Moreover, for large momentum transfers ("hard process"), scattering amplitudes for partons can be calculated perturbatively. This hard process happens at large Q^2, i.e. small distances. The large distance physics (of small Q^2), which are the non-perturbative effects, should, therefore, not affect the hard process significantly, but will happen in space-time far away from the hard interaction. Hence, we expect the cross section to factorize in the hard core-process and a non-perturbative part. This is referred to as factorization.

To establish this intuitive argument on a theoretical level, is highly non trivial. A complete proof is not found yet for all observables [15].

In general, one has to rethink carefully, what exactly is measured in experiments at hadron colliders and what is not measured. A detector measures energy and momentum flow. It always has a finite resolution, thus, particles of very low energy cannot be resolved and two particles, which are collinear to each other, may look like a single particle. Moreover, a detector in a scattering experiment can never cover the whole phase space, but will have blind areas around the beam pipes. Hence, of all the constituents of the hadrons in the initial state we will mostly measure the particles originating from partons that took part in a hard scattering and received a large transverse momentum in this way, while the pieces that did not interact hard, and therefore still move approximately parallel to the beam, mostly will go to the blind areas of our detector. Thus, in an experiment, we never have the full information of the complete final state. Instead, we expect that there are a lot of additional particles, which we did not observe: a group of particles, which are close to the beam pipe, and a group of particles, which are unobserved, since they are soft or collinear to other particles.

Therefore, useful observables should also be approximately insensitive to these particles. Given the factorization of the long and short distance parts and allowing remnants close to beam direction, we do not need to completely describe the initial hadrons, but we only

need to know the chance to find the parton, which participates in the hard scattering, inside the hadron. This simplification will allow factorization in the initial-state. When we look at jets of particles in the final-state, these jets will be approximately insensitive to emissions of additional soft and collinear particles as well as to soft interactions. Hence, we do not need to describe the formation of hadrons in the final state.

Note that factorization is usually not an exact result, but holds up to corrections that are suppressed by inverse powers of the hard momentum scale Q. The form of the factorized cross section is [18]

$$Q^2 \sigma_{phy}(Q, m) = \omega_{\text{SD}}(Q/\mu, \alpha_S(\mu)) \otimes f_{\text{LD}}(\mu, m) + \mathcal{O}(1/Q^p) \,. \tag{2.35}$$

Here Q is the hard scale and m collectively denotes various soft scales like masses. SD abbreviates short distance, LD long distance. $\otimes$ is a convolution. For DIS and hadron-hadron collisions the convolution is in the partonic momentum fraction x, f are the parton distribution functions and ω_{SD} are partonic cross sections. By applying a Mellin transformation, i.e. $\bar{f}(N) = \int_0^1 \mathrm{d}x \, x^{N-1} f(x)$ etc., we get a simple product. The Q and the m dependence are separated in the short and long distance part. The short distance part ω_{SD} does not depend on the soft scales m and is calculable in perturbation theory, while the long distance part f_{LD} depends on the soft scales m and is not calculable perturbatively.

To actually achieve this separation, the new scale μ, the factorization scale, had to be introduced. It arises during the regularization of the integrals of the scattering cross section to remove the divergences from ω_{SD}, which we now calculate in the zero mass limit, and absorb them in f_{LD}. A widely used method of regularization is dimensional regularization, where the scale μ appears due to dimensional reasons. Dimensional regularization was already mentioned during the discussion of UV divergences, but here the ϵ in $d = 4 - 2\epsilon$ is negative. Other regularization prescriptions set a lower bound on the transverse momentum or the virtuality of particles attending the hard scattering.

How can we use eq. (2.35), although f_{LD} is not calculable in perturbation theory? Since f_{LD} factorizes from the short distance part, they are portable from one process to another. The useful fact is that we can determine their dependence on the factorization scale. This follows directly from the Mellin transformed version of eq. (2.35). The physical cross section cannot depend on the factorization scale, since it was introduced arbitrarily to perform the calculation. Hence,

$$\mu \frac{d}{d\mu} \log \bar{\sigma}_{phy}(Q, m) = 0 \,. \tag{2.36}$$

Thus, the μ dependence of the long and short distance part have to cancel:

$$\mu \frac{d}{d\mu} \log \bar{\omega}_{\text{SD}} = -\bar{P}(\alpha_S(\mu)) = -\mu \frac{d}{d\mu} \log \bar{f}_{\text{LD}} \,, \tag{2.37}$$

where $\bar{P}$ can only depend on the common variables of $\bar{\omega}_{\text{SD}}$ and $\bar{f}_{\text{LD}}$, which are $\alpha_S(\mu)$ (or equivalently μ) and the (Mellin transformed) convolution variables. Eq. (2.37) is an evolution equation for $\bar{f}_{\text{LD}}$ with respect to μ. Since we can calculate $\bar{\omega}_{SD}$, taking its derivative we obtain $\bar{P}$. In the next chapter we will determine $\bar{P}$.

3 Parton splitting

The aim of this chapter is to find a simplified form to describe emissions of collinear particles from a core process. This form will be used in the end of this section to describe the evolution of parton density functions. Moreover, it is also the underlying relation of the Monte Carlo parton shower and the resummation, which will be discussed in subsequent chapters.

For hadron-hadron scattering, which is the process we will discuss in subsequent chapters, factorization, as outlined in the last chapter, means more explicitly

$$\sigma_{h_1 h_2 \to Y} = \sum_{a_1, a_2} \int \mathrm{d}x_1 \int \mathrm{d}x_2 f_{a_1/h_1}(x_1, \mu_f) f_{a_2/h_2}(x_2, \mu_f) \hat{\sigma}_{a_1 a_2 \to Y}(x_1, x_2, \mu_f, \mu_r, ...) , \quad (3.1)$$

where $\hat{\sigma}_{a_1 a_2 \to Y}$ is the hard, partonic cross section of the partons a_1 and a_2 to produce the final state Y. $f_{a_i/h_i}(x_i, \mu_f)$ are the parton distribution functions (pdf), which give the "probability" to find parton a_i with a longitudinal momentum fraction x_i and a virtuality absolute value $|t| \leq \mu_f^2$ inside the hadron h_i.

In QCD, f_{a_i/h_i} cannot be calculated perturbatively. Nevertheless, perturbation theory tells us, in which way f depends on μ_f, which is given by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equation [19, 20]. Moreover, f does not depend on the hard, partonic process. Thus, eq. (3.1) has predictive power. Once determined in some process for one scale, we can use f_{a_i/h_i} for any other process and any scale, which is still in the perturbative regime.

To derive the DGLAP equations in the end of this chapter, we will now look at splittings of partons. Since we aim for an expression like eq. (2.35), where the dependence on soft scales like masses has been removed from the perturbatively calculable short distance part, we will also go to the limit of vanishing masses for all partons. As depicted in Figure 3.1, for a given hard partonic matrix element $\mathcal{M}_n$ involving n external partons, we take, for example, the outgoing line of some parton a and attach a vertex, which splits parton a in the two partons b and c. We will find that the contribution of such diagrams to the cross section is not suppressed by a factor of α_S relative to the diagram without the additional splitting, but just by α_S times some potentially large logarithm.

3.1 Final state splitting

We consider the branching of parton a to partons b and c in Figure 3.1a. The partons b and c should be (approximately) on shell, because we want to consider them as external,

(a) outgoing   (b) incoming
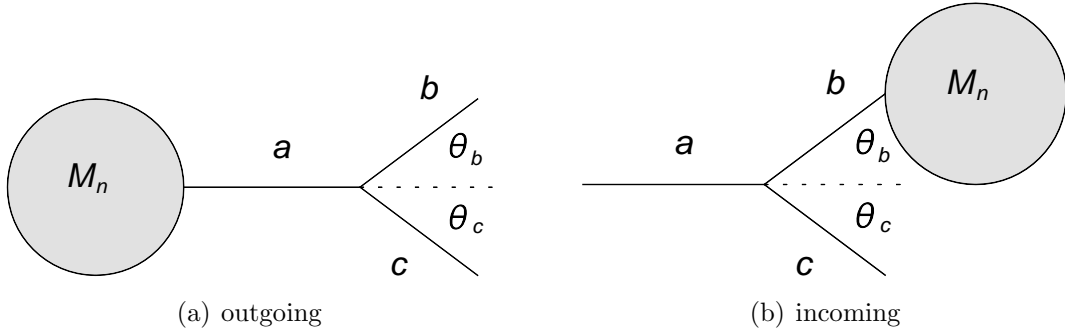
Figure 3.1: An external parton from the hard partonic process which is depicted by the shaded circle is involved in an additional splitting. This parton is (a) the parton a in the final state that splits to partons b and c or (b) the parton b in the initial state that was produced together with parton c in a splitting of parton a.

whereas parton a is virtual and therefore can be off-shell. In fact, it has to be off-shell to produce two non parallel partons - recall that we perform our calculation in the massless limit. Thus, we assume

$$p_b^2, p_c^2 \ll p_a^2 = t > 0 \,. \tag{3.2}$$

The energy fraction, carried by parton b, is

$$z = \frac{E_b}{E_a} = 1 - \frac{E_c}{E_a} \,, \tag{3.3}$$

where the last equality follows from energy conservation at the vertex. The scattering angles of partons b and c with respect to parton a are θ_b and θ_c respectively. Hence, the opening angle between them is $\theta = \theta_b + \theta_c$. For small angles θ and for massless on-shell partons b and c, we have

$$\begin{aligned} t = p_a^2 = (p_b + p_c)^2 &= 2p_b p_c = 2E_b E_c (1 - \cos\theta) \\ &= 2z(1-z)E_a^2(1 - \cos\theta) \approx z(1-z)E_a^2\theta^2 \,. \end{aligned} \tag{3.4}$$

Depending on the choice, we make for the type of partons a, b and c, in QCD there exist three different types of splittings:

- a gluon splitting to two gluons as shown in Figure 3.2(a),

- a gluon splitting to a quark-antiquark pair as shown in Figure 3.2(b),

- a quark emitting a gluon as shown in Figure 3.2(c).

Different choices of quark-flavors and the emission of a gluon from an antiquark are related to these by symmetries.

(a) A gluon splits into two gluons.

(b) A gluon splits into a quark-antiquark pair.

(c) A quark emits a gluon.

Figure 3.2: The three distinct splittings of QCD.

For all three splittings, the calculation is performed in a similar way. As a first step, we express the propagator between the new splitting vertex and the old partonic part by a sum of all physical polarization and color states
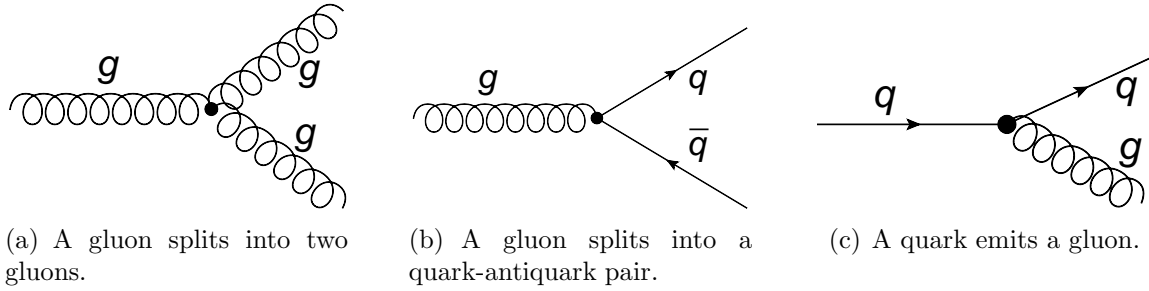
$$\frac{\not{p}}{p^2} = \frac{1}{t} \sum_{s,c} u_{s,c}(p)\bar{u}_{s,c}(p) \quad \text{for a quark,} \tag{3.5}$$

$$\frac{\not{p}}{p^2} = \frac{1}{t} \sum_{s,c} v_{s,c}(p)\bar{v}_{s,c}(p) \quad \text{for an antiquark,} \tag{3.6}$$

$$\frac{1}{p^2}\left(-g_{\mu\nu} + \frac{n_\mu p_\nu + n_\nu p_\mu}{n \cdot p}\right) = \frac{1}{t} \sum_{s,c} \epsilon_\mu^{s,c}(p)(\epsilon_\nu^{s,c}(p))^* \quad \text{for a gluon.} \tag{3.7}$$

In this notation we condensed the spin and color dependence in a single vector.[1] For the gluon $\epsilon \cdot p = 0 = \epsilon \cdot n$ and the sum is only over the transverse polarizations. The sums over s and c extend over all relevant spin and color states, which are, of course, different for the quark and the gluon.

The next step is to split the full matrix element $\mathcal{M}_{n+1}$ with partons b and c in the final state into the old matrix element $\mathcal{M}_n$ with a in the final state and the splitting vertex $\mathcal{M}_{a \to bc}$. To do so, we use the above expressions for the propagator of a. For all explicit choices of momentum, color and helicity of the external particles in the diagrams of $\mathcal{M}_{n+1}$, the conservation of these attributes will project the parton a on a single state. For each explicit choice, we then have [8]

$$|\mathcal{M}_{n+1}|^2 \sim |\mathcal{M}_{a \to bc}|^2 \frac{1}{t^2}|\mathcal{M}_n|^2$$

$$\sim \frac{4g^2}{t}\hat{C}_{a \to bc}F_{a \to bc}(z)|\mathcal{M}_n|^2 \,, \tag{3.8}$$

where C is the color function and, for each choice of polarizations, F is a function of z. One power of t of the squared propagator has been canceled by a power of t in the splitting

---

[1] One can also split them. E.g. for the first relation, one can write $\delta_{ij}\frac{\not{p}_{\alpha\beta}}{p^2} = \frac{1}{t}\sum_{s,c} u_\alpha^s(p)d_i^c\bar{u}_\beta^s(p)(d_j^c)^*$, with the usual Lorentz spinors u and a color vector d obeying $\sum_c d_i^c(d_j^c)^* = \delta_{ij}$.

$a \to bc$	$C_{a \to bc}$	$\langle F \rangle_{a \to bc}$	$\hat{P}_{b \leftarrow a} = C \cdot \langle F \rangle$
$g \to gg$	$3 = \frac{1}{8} f^{ABC} f^{ABC} = C_A$	$\left[\dfrac{1-z}{z} + \dfrac{z}{1-z} + z(1-z) \right]$	$\hat{P}_{g \leftarrow g}(z)$
$g \to q\bar{q}$	$\frac{1}{2} = \frac{1}{8} Tr(t^A t^A) = T_F$	$\left[z^2 + (1-z)^2 \right]$	$\hat{P}_{q \leftarrow g}(z)$
$q \to qg$	$\frac{4}{3} = \frac{1}{3} Tr(t^A t^A) = C_F$	$\left[\dfrac{1+z^2}{1-z} \right]$	$\hat{P}_{q \leftarrow q}(z)$

Table 3.1: The three splittings of QCD. $\hat{P}_{b \leftarrow a} = C \cdot \langle F \rangle$ are the splitting kernels. $C_A = T_A$, C_F and T_F have been defined in eqs. (2.4) and (2.5).

vertex. We assume that, when summing over color and helicity of all other partons, $\mathcal{M}_n$ does not depend on the color and helicity of parton a. Then, summing over final and averaging over initial colors and helicities (as depicted by the bar over $|\mathcal{M}|^2$), gives

$$\overline{|\mathcal{M}_{n+1}|^2} \sim \frac{16\pi\alpha_S}{t} C_{a \to bc} \langle F \rangle_{a \to bc}(z) \overline{|\mathcal{M}_n|^2} \,. \tag{3.9}$$

For the three distinct splittings, the results are given in Table 3.1.
$\hat{P}_{b \leftarrow a} = C \cdot \langle F \rangle$ are the **splitting kernels** containing all information, which will be relevant in the following. The other possible splitting kernels are related to those in Table 3.1 by the relations

$$\hat{P}_{\bar{q} \leftarrow g}(z) = \hat{P}_{q \leftarrow g}(z) \,, \tag{3.10}$$

$$\hat{P}_{g \leftarrow q}(z) = \hat{P}_{g \leftarrow \bar{q}}(z) = \hat{P}_{\bar{q} \leftarrow \bar{q}}(1-z) = \hat{P}_{q \leftarrow q}(1-z) \,. \tag{3.11}$$

They are diagonal in the quark flavors. The splitting kernels are also known to a higher order in perturbation theory than the LO results given above. They can be found e.g. in [8].

The third step is to establish the relation corresponding to eq. (3.9) on cross section level. For this purpose, we have to look at the phase space, which is $\prod_{f \in \text{final}} \frac{\mathrm{d}^3 p_f}{(2\pi)^3 2E_f}$. There, due to the splitting, we have to replace $\frac{\mathrm{d}^3 p_a}{(2\pi)^3 2E_a}$ by $\frac{\mathrm{d}^3 p_b}{(2\pi)^3 2E_b} \frac{\mathrm{d}^3 p_c}{(2\pi)^3 2E_c}$. Using $p_c = p_a - p_b$, we can write

$$\frac{\mathrm{d}^3 p_c}{(2\pi)^3 2E_c} = \frac{\mathrm{d}^3 p_a}{(2\pi)^3 2E_a} \cdot \frac{1}{1-z} \tag{3.12}$$

to get back the old phase space times the integration over the phase space of b. The latter can be rewritten in the limit of a small scattering angle by a variable transformation from (θ_b, E_b) to (t, z), which are defined in eqs. (3.4) and (3.3), as

$$\frac{\mathrm{d}^3 p_b}{(2\pi)^3 2E_b} = \frac{1}{2(2\pi)^3} E_b \mathrm{d}E_b \theta_b \mathrm{d}\theta_b \mathrm{d}\phi = \frac{1}{2(2\pi)^3}(1-z)\mathrm{d}z\mathrm{d}t\mathrm{d}\phi \,. \tag{3.13}$$

Therefore, we can write the new phase space in terms of the old one, times three additional integrations,

$$d\Phi_{n+1} = d\Phi_n \frac{1}{4(2\pi)^3} dt\, dz\, d\phi\,. \tag{3.14}$$

Furthermore, using eq. (3.9), we find for the cross section

$$d\sigma_{n+1} = d\sigma_n \frac{dt}{t} dz \frac{d\phi}{2\pi} \frac{\alpha_S}{2\pi} C\langle F\rangle\,. \tag{3.15}$$

Integrating over ϕ and using the splitting kernel $\hat{P} = C\langle F\rangle$, yields

$$\boxed{d\sigma_{n+1} = d\sigma_n \frac{dt}{t} dz \frac{\alpha_S}{2\pi} \hat{P}_{b\leftarrow a}(z)\,.} \tag{3.16}$$

Due to the factors $\frac{1}{t}$ and $\hat{P}$, which can contain a term with z or $1-z$ in the denominator, the splitting is dominated by phase space configurations, in which the virtuality of a is small or partons b or c carry a small energy fraction. These are the collinear and soft configurations.

3.2 Initial state splitting

The results and discussion is similar for splittings in the initial state. In the process under consideration, we now create the parton b, which has been in the initial state of $\mathcal{M}_n$, from a splitting of parton a to partons b and c, as shown in Figure 3.1b. The propagator of b can be split as in eqs. (3.5) - (3.7). On matrix element level, we then find the same kind of factorization as in eq. (3.9) with the same splitting kernels at LO. In two places, we encounter differences to the final state splitting: First, the definition of t, which is now given by

$$-t = -p_b^2 \approx E_a E_c \theta_c^2 \approx z^2(1-z)E_a^2\theta \quad \text{with} \tag{3.17}$$
$$-p_b^2 \gg |p_a^2|, |p_c^2|\,, \tag{3.18}$$

and, second, the phase space integration. This time, the phase space $d\Phi_n$ is trivially contained in $d\Phi_{n+1} = d\Phi_n \frac{d^3p_c}{(2\pi)^3 2E_c}$ and we write

$$\frac{d^3p_c}{(2\pi)^3 2E_c} = \frac{1}{4(2\pi)^3} dt \frac{dz}{z} d\phi\,. \tag{3.19}$$

Moreover, to get back $d\sigma_n$, we have to replace $\frac{1}{E_b}$ in the initial state flux factor by $\frac{1}{E_a} = \frac{z}{E_b}$. Hence, we again find eq. (3.16)

$$d\sigma_{n+1} = d\sigma_n \frac{dt}{t} dz \frac{\alpha_S}{2\pi} \hat{P}_{b\leftarrow a}(z)\,. \tag{3.20}$$

3.3 Large logarithms

For small splitting angles and for the approximation in eq. (3.18) or (3.2), we can replace the t integration in eq. (3.20) or (3.16), respectively, by an integration over the transverse momentum $\vec{p}_\perp^2$ of the splitting products. We than have

$$\frac{\mathrm{d}\sigma_{n+1}}{\mathrm{d}\sigma_n} = \frac{\mathrm{d}\vec{p}_\perp^2}{\vec{p}_\perp^2}\mathrm{d}z\frac{\alpha_S}{2\pi}\hat{P}_{b\leftarrow a}(z)\,. \tag{3.21}$$

There are potential divergences, when we try to perform the $\vec{p}_\perp^2$ and z integrations. Obviously, the $\vec{p}_\perp^2$ integral diverges at $\vec{p}_\perp^2 \to 0$. Moreover, the splitting kernels given in Table 3.1 contain terms as $1/z$ and $1/(1-z)$. After integration the latter two will lead to $\log(\frac{z_+}{z_-}) = -\log(\frac{1-z_+}{1-z_-})$ with z_+ and z_- being the upper and lower bound of the z integration, respectively. This, too, diverges, for $z_- \to 0$ and $z_+ \to 1$.

We want to discuss in this section, how we can regulate these divergences by introducing an explicit lower cut-off on the $\vec{p}_\perp^2$ integration and what the most relevant contributions of the regulated terms will be. We will also consider multiple successive splittings.

For a given value of $\vec{p}_\perp^2$, the energy of the two created particles is limited from below by $|\vec{p}_\perp| \leq E$. Thus,

$$z_- = 1 - z_+ = \frac{|\vec{p}_\perp|}{E_a} \tag{3.22}$$

with the energy E_a of the splitting particle, which in turn is limited by the total energy of the scattering process. Here, the first equality follows from energy conservation and the corresponding limit for the other emitted particle. Energy conservation also dictates

$$|\vec{p}_\perp| \leq |\vec{p}_{\perp,\mathrm{max}}| = \frac{1}{2}E_a\,. \tag{3.23}$$

Thus, the z integration gives

$$\int_{z_-}^{z_+} \mathrm{d}z\frac{\alpha_S}{2\pi}\hat{P}_{b\leftarrow a}(z) = \alpha_S 2A \log\left(\frac{\vec{p}_\perp^2}{\vec{p}_{\perp,\mathrm{max}}^2}\right) + \alpha_S B + \text{subleading} \tag{3.24}$$

with some constants A and B, which will be determined in section 5 for a similar case. "Subleading" is anything that vanishes if $\vec{p}_\perp^2 \to 0$. This also includes the contributions from $\log(z_+) = \log(1 - z_-) = \log(1 - \frac{|\vec{p}_\perp|}{2|\vec{p}_{\perp,\mathrm{max}}|})$, which is limited by $\log\frac{1}{2}$ and vanishes as $|\vec{p}_\perp|$ does.

To regularize the $\vec{p}_\perp^2$ integration, we introduce a lower cut-off $\vec{p}_{\perp,0}^2$ to control the divergence there. With the help of the general result

$$\int_{x_0}^{x_m} \frac{\mathrm{d}x}{x} \log^n x = \int_{\log(x_0)}^{\log(x_m)} \mathrm{d}(\log x) \log^n x = \frac{1}{n+1}\log^{n+1}\frac{x_m}{x_0}\,, \quad n \geq 1\,, \tag{3.25}$$
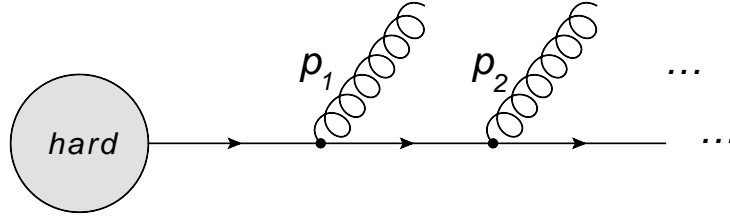
Figure 3.3: A quark, which took part in a hard interaction, emits multiple gluons. The dots refer to further emissions.

we then find

$$\int_{\vec{p}_{\perp,0}^2}^{\vec{p}_{\perp,\text{max}}^2} \frac{\mathrm{d}\vec{p}_\perp^2}{\vec{p}_\perp^2} = \log\left(\frac{\vec{p}_{\perp,\text{max}}^2}{\vec{p}_{\perp,0}^2}\right) \equiv L \text{ and} \tag{3.26}$$

$$\int_{\vec{p}_{\perp,0}^2}^{\vec{p}_{\perp,\text{max}}^2} \frac{\mathrm{d}\vec{p}_\perp^2}{\vec{p}_\perp^2} \log\left(\frac{\vec{p}_\perp^2}{\vec{p}_{\perp,\text{max}}^2}\right) = \frac{1}{2}L^2 \,. \tag{3.27}$$

Hence, for one splitting, the strongest enhanced terms are

$$A\alpha_S L^2 \quad \text{and} \quad B\alpha_S L \,, \tag{3.28}$$

i.e. a leading logarithmic (LL) term, which has two powers of the large logarithm L for a single power of α_S and a next-to-leading logarithmic (NLL) term, which has one power of the large logarithm for a power of α_S. From which regions of phase space did the logarithms arise? The relevant contribution to both the LL and NLL terms was from the collinear region, where $1/\vec{p}_\perp^2$ gets large and produces a logarithm in the integration. The z integration just led to a constant for the NLL term, but for the LL term it already produced a power of the logarithm $\log(\vec{p}_\perp^2/\vec{p}_{\perp,\text{max}}^2)$, as we see from eq. (3.24). This logarithm was due to the terms $1/z$ and $1/(1-z)$ in the splitting kernels, which get large if one splitting product becomes soft, i.e. $z \to 0,1$. Hence, the LL term corresponds to splittings, which are both, collinear and soft while the NLL terms corresponds to splittings, which are just collinear.

Multiple splittings

For multiple splittings, which are ordered in $\vec{p}_\perp^2$, we can apply eq. (3.20) or (3.16) iteratively. By ordered in $\vec{p}_\perp^2$ we mean $\vec{p}_{\perp,1}^2 > \vec{p}_{\perp,2}^2 > ...$, where splitting "1" happens closest to the hard core-process, as depicted in Figure 3.3. The $\vec{p}_{\perp,n}^2$ are with respect to the direction of

the particle entering the nth splitting. For two emissions, we find

$$
\int_{\vec{p}_{\perp,0}^2}^{\vec{p}_{\perp,\max}^2} \frac{\mathrm{d}\vec{p}_{\perp,1}^2}{\vec{p}_{\perp,1}^2} \alpha_S \left(2A \log \frac{\vec{p}_{\perp,1}^2}{\vec{p}_{\perp,\max}^2} + B \right) \int_{\vec{p}_{\perp,0}^2}^{\vec{p}_{\perp,1}^2} \frac{\mathrm{d}\vec{p}_{\perp,2}^2}{\vec{p}_{\perp,2}^2} \alpha_S \left(2A' \log \frac{\vec{p}_{\perp,2}^2}{\vec{p}_{\perp,1}^2} + B' \right)
$$

$$
= \int_{\vec{p}_{\perp,0}^2}^{\vec{p}_{\perp,\max}^2} \frac{\mathrm{d}\vec{p}_{\perp,1}^2}{\vec{p}_{\perp,1}^2} \alpha_S \left(2A \log \frac{\vec{p}_{\perp,1}^2}{\vec{p}_{\perp,\max}^2} + B \right) \alpha_S \left(A' \log^2 \frac{\vec{p}_{\perp,1}^2}{\vec{p}_{\perp,0}^2} + B \log \frac{\vec{p}_{\perp,1}^2}{\vec{p}_{\perp,0}^2} \right) \qquad (3.29)
$$

$$
= \alpha_S^2 C_2 \log^4 \left(\frac{\vec{p}_{\perp,\max}^2}{\vec{p}_{\perp,0}^2} \right) + \alpha_S^2 D_2 \log^3 \left(\frac{\vec{p}_{\perp,\max}^2}{\vec{p}_{\perp,0}^2} \right) + \mathcal{O}(\text{NNLL}) \,.
$$

In the last step, we used $\log(\vec{p}_{\perp,1}^2/\vec{p}_{\perp,0}^2) = \log(\vec{p}_{\perp,1}^2/\vec{p}_{\perp,\max}^2) - \log(\vec{p}_{\perp,\max}^2/\vec{p}_{\perp,0}^2)$. Accordingly, for n splittings ordered in $\vec{p}_\perp^2$, we find

$$
C_n \alpha_S^n \log^{2n} \frac{\vec{p}_{\perp,0}^2}{\vec{p}_{\perp,\max}^2} + D_n \alpha_S^n \log^{2n-1} \frac{\vec{p}_{\perp,0}^2}{\vec{p}_{\perp,\max}^2} + \mathcal{O}(\text{NNLL}) \,. \qquad (3.30)
$$

Therefore, at order n of perturbation theory, there will arise a LL term $\sim \alpha_S^n L^{2n}$ and a NLL term $\sim \alpha_S^n L^{2n-1}$ of the large logarithmic term $L = \log(\vec{p}_{\perp,0}^2/\vec{p}_{\perp,\max}^2)$. For the LL term, all emissions have to be collinear and soft, while for the NLL term, one emission can be just collinear. From eq. (3.30) we also see, that for a series of collinear and soft emissions, the effective expansion is not anymore in terms of $\alpha_S(\mu_r)$, but in terms of $\alpha_S(\mu_r)$ multiplied by large logarithms. These slow down or even spoil the convergence of the perturbative expansion. Hence, we should try take the effect of the splittings enhanced by large logarithms into account to all orders. Methods to do so will be presented in chapters 4 and 6.

The logarithms represent the collinear and infrared divergences of QCD. In an analogous form, they are also present in Quantum Electro Dynamics (QED), as can be seen from our discussion for QCD, when ignoring color and the triple-gauge-boson-vertex and with $\alpha_S \to \alpha_{em}$. In QED, these divergences can be seen as a manifestation of the correspondence principle. To this end, let us recall, what we expect in classical electrodynamics for a fast moving charge, which takes part in a scattering process, i.e. which is accelerated. We expect it to radiate. Since the charge moves fast, the radiation should be mostly collinear. The best quantum approximation for a classical field are almost infinitively many, very soft photons. Thus, finding our collinear ($\vec{p}_\perp^2 \to 0$) and infrared ($z \to 0, 1$) singularities in the corresponding quantum theory (eq. (3.16)) is maybe even reassuring.

Let us now look at two emissions disordered in $\vec{p}_\perp^2$, i.e. this time we have $\vec{p}_{\perp,1}^2 < \vec{p}_{\perp,2}^2$. The virtuality of the parton, propagating to its first splitting, is therefore not of order $\vec{p}_{\perp,1}^2$, but of order $\vec{p}_{\perp,2}^2$. As discussed along with eq. (3.8), it is the virtuality, that enters the nominator. Given that an analogous relation to eq. (3.20)/(3.16) still holds, we then have

for the inner two integral, which are now for particle "1",

$$\int_{\vec{p}_{\perp,0}^2}^{\vec{p}_{\perp,2}^2} \frac{\mathrm{d}\vec{p}_{\perp,1}^2}{\vec{p}_{\perp,2}^2} \int_{z_-(\vec{p}_{\perp,2}^2)}^{z_+(\vec{p}_{\perp,2}^2)} \mathrm{d}z_1 \frac{\alpha_S}{2\pi} \hat{P}(z_1) \sim \int_{\vec{p}_{\perp,0}^2}^{\vec{p}_{\perp,2}^2} \alpha_S A \frac{\vec{p}_{\perp,1}^2}{\vec{p}_{\perp,2}^2} + \alpha_S B$$

$$= \alpha_S A \left(-\frac{\vec{p}_{\perp,0}^2}{\vec{p}_{\perp,2}^2} \log \frac{\vec{p}_{\perp,0}^2}{\vec{p}_{\perp,2}^2} - \frac{\vec{p}_{\perp,2}^2 - \vec{p}_{\perp,0}^2}{\vec{p}_{\perp,2}^2} \right) + \alpha_S B \frac{\vec{p}_{\perp,2}^2 - \vec{p}_{\perp,0}^2}{\vec{p}_{\perp,2}^2}. \tag{3.31}$$

If we perform additionally the $\vec{p}_{\perp,2}^2$ and z_2 integration, we find that non of this terms gives a LL or NLL contribution. As we see from eq. (3.29), this would require at least three powers of the large logarithm $\log(\vec{p}_{\perp,\mathrm{max}}^2/\vec{p}_{\perp,0}^2)$, but both $|\vec{p}_{\perp,0}^2/\vec{p}_{\perp,2}^2 \log(\vec{p}_{\perp,0}^2/\vec{p}_{\perp,2}^2)|$ and $|(\vec{p}_{\perp,2}^2 - \vec{p}_{\perp,0}^2)/\vec{p}_{\perp,2}^2|$ are smaller than 1 in the whole integration range of $\vec{p}_{\perp,2}^2$. Hence, the largest contribution, which can arise, is $\alpha_S L^2$, which is NNLL.

Therefore, we find that at NLL accuracy, just emissions ordered in $\vec{p}_\perp^2$ need to be considered. This in turn motivates the assumptions in eq. (3.18) and (3.2).

3.4 Parton distribution functions

When we try to evaluate the successive splittings perturbatively, this will work for a notably amount of them, but when we reach a point in our cascade, where the involved transverse momenta get too small or, equivalently, the virtuality of the internal parton comes close to zero, we cannot use perturbative methods anymore. Whatever their effect might be, due to the effective ordering in t, this kind of splittings will effectively just appear after (for final state splitting) or before (for initial state splittings) the splitting cascade, which we can describe perturbatively.

Thus, for an initial parton, we can consider these splitting as some approximately continuous process, which evolves a parton of very low virtuality to a parton of large (negative) virtuality, which participates in a hard scattering. This parton will just carry a fraction z of the energy of the parton it originates from, because the other emitted partons will carry the rest of that energy. Moreover, since there are also splittings, where the type of the parton changes (like $g \to q\bar{q}$ or $q \to gq$), the type of parton may have changed. Formulated differently, when we probe a parton a at a given virtuality scale t, for each parton type b and energy fraction z, there is a chance $\tilde{f}_{b/a}(z,t)$ to scatter with a parton of type b carrying the fraction z of the energy of parton a. Probing at a scale t_p, i.e. scattering at sufficiently large energies, we resolve all partons of virtuality $-t \leq -t_p$. This is, because a high virtuality corresponds to a small lifetime (before being involved in another interaction) and at a given energy we can just resolve particles, which have more than some minimal lifetime. Since the partons b are created by successive splittings staring with parton a, we can use what we learned about splittings in the proceeding part to learn something about the distributions $\tilde{f}$.

Given the distributions $f_g(z,t)$, $f_{q_f}(z,t)$ and $f_{\bar{q}_f}(z,t)$ at some scale t, the corresponding distributions at the scale $t - \Delta t$ (keep in mind $t < 0$) should follow from the former by

taking an additional splitting into account. As argued above, we will still see partons with virtuality $-t' \leq -t$, but now we additionally see the partons with virtuality $t' \in [t - \Delta t, t]$, which are generated by splittings of the other partons. Thus,

$$
f_b(z, t - \Delta t) = f_b(z, t) + \int_0^1 dx \int_0^1 dz' \int_{t-\Delta t}^t \frac{dt'}{t'} \frac{\alpha_S}{2\pi} \sum_a \hat{P}_{b \leftarrow a}(z') f_a(x, t) \delta(z - z'x)
$$

$$
- \int_0^1 dz' \int_{t-\Delta t}^t \frac{dt'}{t'} \frac{\alpha_S}{2\pi} \sum_a \left| \hat{P}_{a \leftarrow b}(z') f_b(z, t) \right|_{\text{no } 2\times} \tag{3.32}
$$

$$
= f_b(z, t) + \sum_a \frac{\Delta t}{t} \int_0^1 \frac{dz'}{z'} \frac{\alpha_S}{2\pi} \hat{P}_{b \leftarrow a}(z') f_a(\frac{z}{z'}, t) - \sum_a \left| \frac{\Delta t}{t} \right|_{\text{no } 2\times} \int_0^1 dz' \frac{\alpha_S}{2\pi} \hat{P}_{a \leftarrow b}(z') f_b(z, t) \,.
$$

The second term accounts for the new partons of momentum fraction z and high (negative) virtuality $t' \in [t - \Delta t, t]$. The first and the third term instead account for the partons b of momentum fraction z and virtuality $-t' \leq -t$, which are the old ones $f_b(z, t)$ (first term) minus the number of partons that are gone, since they split (third term). To not count the removal of a single parton twice, the sum in the third term is restricted to a subset of possible partons a. For $b = q\,(\bar{q})\,[g]$, we choose $a = q\,(\bar{q})\,[g, q]$, since e.g. for $b = q$ also including $a = g$, in addition to $a = q$, would mean to count the splitting of $a = q \to qg$ twice. The sum in the second term extends over all partons a. Here, the splitting $\hat{P}_{g \to g}$ needs to be included two times, since the splitting $g \to gg$ produces two gluons (this splitting is symmetric under $z \to 1 - z$, thus, we can just include the same term twice).

When we look at eq. (3.32), we see the singularity at $t' \to 0$ is avoided, since the integral is bounded from below. The non-perturbative splittings at low virtuality, about which we were concerned, are absorbed in f. Due to energy conservation $f_b(\tilde{z}, t)$ vanishes for $\tilde{z} > 1$. Hence, in the second term there is no contribution from $z' < z$ and the singularities of the splitting kernels at $z' \to 0$ are avoided. The only possible pole left in this term is for $z' \to 1$. Looking at the splitting kernels in table 3.1 and how the other splitting kernels are related to them (eqns. (3.10,3.11)), we see that just the diagonal splitting kernels, i.e. $\hat{P}_{q \leftarrow q}$, $\hat{P}_{\bar{q} \leftarrow \bar{q}}$ and $\hat{P}_{g \leftarrow g}$, have a pole at $z' \to 1$. All of these diagonal splittings also appear in the third term of eq. (3.32), which also has poles like this, but with the opposite sign. Since in the limit $z' \to 1$ the integrands for diagonal splittings of the second and third term are equal[2], the poles will cancel. To make this more explicit, we combine the second and the third term:

$$
\frac{\Delta t}{t} \int_0^1 \frac{dz'}{z'} \frac{\alpha_S}{2\pi} \hat{P}_{a \leftarrow a}(z') f_a(\frac{z}{z'}, t) - \frac{\Delta t}{t} \int_0^1 dz' \frac{\alpha_S}{2\pi} \hat{P}_{a \leftarrow a}(z') f_a(z, t)
$$

$$
\equiv \frac{\Delta t}{t} \int_0^1 \frac{dz'}{z'} \frac{\alpha_S}{2\pi} \left(\hat{P}_{a \leftarrow a}(z') \right)_+ f_a(\frac{z}{z'}, t) \,, \tag{3.33}
$$

where for any smooth functions $f(z')$ and $g(z')$ with $z' \in [0, 1[$, f smooth in $z' = 1$ and g

---

[2]The gluon term needs a closer look, which is done after eq. (3.38).

with a simple pole there, we did define the plus prescription:

$$\int_0^1 dz' \, (g(z'))_+ \, f(z') = \int_0^1 dz' \, (f(z') - f(1)) \, g(z') \tag{3.34}$$

or written differently

$$(g(z'))_+ = \lim_{\epsilon \to 0} \left(\Theta(1 - z' - \epsilon)g(z') - \delta(1 - z') \int_0^{1-\epsilon} dz' g(z') \right) . \tag{3.35}$$

Form eq. (3.34) it is obvious that for the given functions the integral is finite. We now define the new splitting functions (without a hat) as

$$P_{\bar{q}\leftarrow\bar{q}}(z) = P_{q\leftarrow q}(z) = \left(\hat{P}_{q\leftarrow q}(z) \right)_+ = C_F \left[\frac{1 + z^2}{(1 - z)_+} + \frac{3}{2}\delta(1 - z) \right] , \tag{3.36}$$

$$P_{g\leftarrow g}(z) = 2 \left(\hat{P}_{g\leftarrow g}(z) \right)_+ - n_f \delta(1 - z) \int_0^1 dy \hat{P}_{q\leftarrow g}(y)$$

$$= 2C_A \left[\frac{z}{(1 - z)_+} + \frac{1 - z}{z} + z(1 - z) \right] + \frac{1}{6}[11C_A - 4n_f T_F]\delta(1 - z) , \tag{3.37}$$

$$P_{b\leftarrow a}(z) = \hat{P}_{b\leftarrow a}(z) \quad \text{if } a \neq b . \tag{3.38}$$

Writing the '+' index just on the divergent nominator and not on the whole term makes a difference at $z \to 1$, where an additional term with a δ-function appears, as can be seen from eq. (3.35), where it corresponds to a change of $g(z')$. Some words on $P_{g\leftarrow g}$ are in place: In the second term of eq. (3.32) it appeared twice, since we had to consider both gluons in the process $g \to gg$. We used the symmetry $\hat{P}_{g\leftarrow g}(1 - z) = \hat{P}_{g\leftarrow g}(z)$ to combine both contributions resulting in the factor 2 in eq. (3.37). In the third term of eq. (3.32) the integral ranges from 0 to 1. The gluon splitting kernel has a pole at 1 and at 0. For the contribution of the pole at 0 we again used the symmetry of the gluon splitting kernel to map it on a pole at 1. Hence, we end up with the correct terms to define $2 \left(\hat{P}_{g\leftarrow g}(z) \right)_+$ in eq. (3.37). We also included the term involving $\int_0^1 dy \hat{P}_{q\leftarrow g}(y)$ in the gluon function times the number of active quark flavors n_f. Now we can rewrite eq. (3.32) as

$$\frac{t}{\Delta t} [f_b(z, t - \Delta t) - f_b(z, t)] = \sum_a \int_0^1 \frac{dz'}{z'} \frac{\alpha_S}{2\pi} P_{b\leftarrow a}(z') f_a\left(\frac{z}{z'}, t\right) , \tag{3.39}$$

where now the sum is over all partons ($P_{g\leftarrow g}$ not counted twice anymore). In the limit $\Delta t \to 0$ this is

$$\boxed{t\frac{df_b(z, t)}{dt} = \sum_a \int_0^1 \frac{dz'}{z'} \frac{\alpha_S}{2\pi} P_{b\leftarrow a}(z') f_a\left(\frac{z}{z'}, t\right)} \tag{3.40}$$

$$\equiv \sum_a \frac{\alpha_S}{2\pi} \left(P_{b\leftarrow a} \otimes f_a \right)(z, t) .$$

3 Parton splitting

These are the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [19,20]. They govern the evolution of the parton density functions (pdf) $f_b(z, t)$. Once determined at one scale t, the pdfs can be found at any scale, that still justifies the perturbative approach. Moreover, since we did not assume anything about the hard process, in which the parton attends, the pdfs and their evolution do not depend on it. All that is set by the hard process are the parton types and the scale t at which the pdfs are evaluated.

The pdfs will, of course, depend on the type of hadron h from which the parton originates. Before attending in a hard scattering process, the parton is involved in many interactions with the other partons of the hadron. To describe all of them properly in perturbation theory, is not possible, especially because many of them are soft and thus enhanced. However, we noted before, that the emissions are effectively ordered in p_T, such that the virtuality $|t|$ of the parton will increase during its evolution. At some sufficiently large scale $|t| = \mu_0$ we can start to describe the evolution of the pdfs perturbatively. All that enters from the preceding evolution is the distribution $f_{p/h}(z, \mu_0)$ at the scale μ_0, which cannot be determined perturbatively, but must be measured experimentally. Nevertheless, the evolution of the distribution can be determined from the DGLAP equations. This works for both directions, to larger or smaller values of μ_0, as long as we do not enter a non-perturbative region. A commonly used set of pdfs for the proton is shown in Figure 3.4 for two different scales. One sees that the pdfs increase rapidly for small values of the momentum fraction x; large values of x are usually just carried by the valence quarks, which define the quantum numbers of the hadron. As one increases the scale μ, the chance to observe a gluon increases.

If we collide two hadrons, any of their partons may interact with each other. Therefore, the scattering cross section of two hadrons to some final state Y is[3]

$$
\begin{aligned}
\sigma_{h_1 h_2 \to Y}(s, ...) = \sum_{a_1 a_2} \int \mathrm{d}x_1 \int \mathrm{d}x_2 f_{a_1/h_1}(x_1, \mu_f) f_{a_2/h_2}(x_2, \mu_f) \hat{\sigma}_{a_1 a_2 \to Y}(x_1, x_2, ..., \mu_f, \mu_r) \\
+ \mathcal{O}(1/Q) \,,
\end{aligned}
\tag{3.41}
$$

as already stated in eq. (3.1). The pdfs will appear in all calculations of cross sections in the following chapters. With the choice μ_f, we can steer, where the corrections due to splittings up to the hard scale are taken into account. For a small choice of scale $\mu_f^2 \ll p_{T,\mathrm{hard}}^2$ we should include corrections from splittings in $\hat{\sigma}$. With a choice $\mu_f^2 \sim p_{T,\mathrm{hard}}^2$ these effects are contained in the pdfs. Hence, this will be our usual choice.

Factorization is actually not an exact result. In our discussion for example we made several approximations. In addition to the ordering in p_T, which was a next-to-leading logarithmic approximation, we also ignored the masses of the partons with respect to the hard scale. Corrections to factorization are suppressed by inverse powers of the hard scale. If this scale is sufficiently high, we can ignore these corrections.

Connecting back to the discussion in the end of chapter 2.2, we find that eq. (3.41) can be written in the form of eq. (2.35): It is a convolution in the partonic momentum fractions

---

[3]We assume that only one pair of partons is involved in the hard scattering process under consideration.
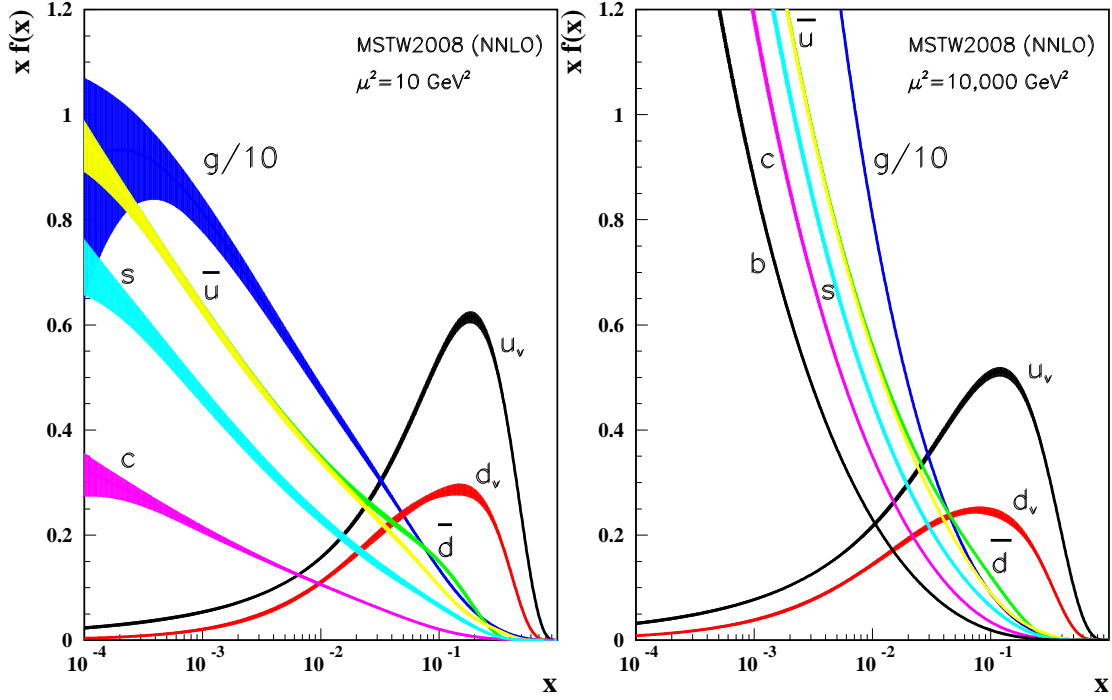
Figure 3.4: The parton distribution functions of the proton multiplied by the energy fraction x for two different scales μ using MSTW2008 NNLO parameterization. Figure extracted from [4].

x, which can be written as

$$\sigma_{h_1 h_2 \to Y}(s, ...) = \sum_{a_1 a_2} (f_{a_1/h_1} \otimes f_{a_2/h_2} \otimes \hat{\sigma}_{a_1 a_2 \to Y})(\mu_f, \mu_r, s, ...) + \mathcal{O}(1/Q). \qquad (3.42)$$

Moreover, the evolution of f, eq. (3.40), is after a Mellin transformation of the form of eq. (2.37). We also argued that by the choice of μ_f we can decide, in which part ($\hat{\sigma}$ or f) we account for the splittings of high virtuality. Hence, we found what we proposed in chapter 2.2.

4 Monte Carlo parton shower

We have seen in section 3.3 that the contribution of higher order terms is logarithmically enhanced if partons in the final state are collinear to each other or soft. Since in this limit these logarithms become large, we want to account for the dominant effect of soft and collinear emissions to all orders in perturbation theory. A method to achieve this is the Monte Carlo parton shower, which will be discussed in this chapter. It allows for in principle arbitrarily many particles in the final state. Emissions of additional particles from a hard core-event are generated with a Monte Carlo method, until a stop criterion is reached. Thereby, the four momenta of these additional particles are generated with a probability distribution, which is in accordance with eq. (3.16). In this way the leading logarithmic terms can be accounted for to all orders.

In our discussion, we will first introduce the Sudakov form factor, which is the central ingredient of all parton showers. Then, we will discuss, how it can be used to generate successive splittings. Finally, we will introduce the Monte Carlo parton shower program SHERPA, whose logarithmic accuracy we will discuss in chapter 5 and which we will study numerically in chapter 8.

4.1 General ideas

Sudakov form factor

Let us start this chapter along the lines of [8, 21] by introducing another formulation of initial state splitting. To this end, we introduce the Sudakov form factor as

$$\Delta_b(t) \equiv \exp\left[-\sum_a \int_{t_0}^t \frac{\mathrm{d}t}{t} \int \mathrm{d}z \frac{\alpha_S}{2\pi} \hat{P}_{a\leftarrow b}(z)\right], \tag{4.1}$$

where $\hat{P}$ are the unregularized splitting kernels and the sum is over all allowed splitting processes. With the Sudakov, we can rewrite the $\Delta t \to 0$ limit of eq. (3.32) as

$$t\frac{\mathrm{d}f_b(z,t)}{\mathrm{d}t} = \sum_a \int_0^1 \frac{\mathrm{d}z'}{z'} \frac{\alpha_s}{2\pi} \hat{P}_{b\leftarrow a}(z') f_a(\frac{z}{z'},t) - \frac{f_b(z,t)}{\Delta_b(t)} t\frac{\mathrm{d}\Delta_b(t)}{\mathrm{d}t}. \tag{4.2}$$

Integrating this equation, we get an expression for $f_b(x,t)$ in terms of the initial parton distributions $f_b(x,t_0)$:

$$f_b(z,t) = \Delta_b(t) f_b(z,t_0) + \int_{t_0}^t \frac{\mathrm{d}t'}{t'} \frac{\Delta_b(t)}{\Delta_b(t')} \int \frac{\mathrm{d}z'}{z'} \frac{\alpha_s}{2\pi} \hat{P}_{b\leftarrow a}(z') f_a(\frac{z}{z'},t). \tag{4.3}$$

The structure of this equation together with our discussion about pdfs in chapter 3.4 suggests the physical interpretation of the Sudakov form factor $\Delta_b(t)$ to be the probability for a parton b to evolve from t_0 to t without resolvable branching. The factor $\frac{\Delta_b(t)}{\Delta_b(t')}$ represents the probability to evolve from t' to t without resolvable branching.

The interpretation of the Sudakov as a non-branching probability can also be motivated in another way. We recall that the Poisson distribution

$$\mathcal{P}(n, \lambda) = \frac{\lambda^n e^{-\lambda}}{n!} \tag{4.4}$$

gives the probability that an event with expected rate λ is observed n times. Thus, the chance to not observe any occurrence is $\mathcal{P}(0, \lambda) = e^{-\lambda}$. This is exactly the form of the Sudakov with $\lambda = \sum_b \int_{t_0}^t \frac{dt'}{t'} \int dz \frac{\alpha_S}{2\pi} \hat{P}_{b\leftarrow a}(z)$. To better understand this term we rewrite eq. (3.16) as

$$\frac{d\sigma_{n+1}}{d\sigma_n} = \frac{dt'}{t'} dz \frac{\alpha_S}{2\pi} \hat{P}_{b\leftarrow a}(z). \tag{4.5}$$

When we integrate this for all $t' \in [t_0, t]$ and $z \in [0, 1]$ and sum over all splitting processes, the right hand side gives λ. As we see from the equation's left hand side, this describes the rate of any splitting to occur. Hence, we again arrive at the interpretation of the Sudakov as a non-splitting probability. We assume here that the probabilistic interpretation of the splitting is valid and that the chosen parameterization of the splitting is a sufficient approximation in the whole integration range.

The integrals in eq. (4.1) involve the unregularized splitting kernels $\hat{P}_{b\leftarrow a}$. Since some of them have divergences at $z = 0$ or $z = 1$, we have to take care of these by e.g. introducing an infrared cut-off $\epsilon(t)$ on the z-integration. In this way the very soft splittings, which we will classify as unresolvable, do not appear in the exponent. Similarly, the splittings with $t < t_0$, which are very collinear and therefore are considered as unresolvable, do not appear in the Sudakov. Although we did not show this, the Sudakov form factor does not just sum enhanced real contributions (parton emissions), but also virtual contributions (parton loops) to all orders. In fact, the coupling constant α_S in the above and following equations is understood to be evaluated at a scale of order $\vec{p}_\perp^2$, e.g. in the $\overline{MS}$ scheme to second order. The virtual corrections affect the non-branching probability. Unitarity dictates that the sum of branching and non-branching probability must be unity. Hence, the divergences for the unresolved real and the virtual contributions have to cancel and our discussion above is sensible.

Generating splittings

The Sudakov form factors are essential for Monte Carlo parton showers. The latter are tools that aim to include the leading effects of soft and collinear emissions to all orders in the calculation of cross sections. Given a hard core-process with n final-state particles the

parton shower generates for each of them successive splittings. In each of these splittings, a given parton reduces its virtuality t and its momentum fraction x. In this way it runs through a series of momentum-virtuality points (x_i, t_i). From the point (x_i, t_i) the following point (x_{i+1}, t_{i+1}) can be generated by a Markov process. We choose a random number r_t, which is equaly distributed in $[0, 1]$. Then, for given t_i, we get t_{i+1} as the solution of

$$\frac{\Delta_a(t_i)}{\Delta_a(t_{i+1})} = r_t \qquad \text{with } t_i > t_{i+1} > t_0 > 0 \,. \tag{4.6}$$

For the gluon, several kinds of splittings are possible. We can choose a splitting $a \to b$ with the probability $\int_{\epsilon(t)}^{1-\epsilon(t)} dz \frac{\alpha_s}{2\pi} \hat{P}_{b\leftarrow a}(z) / \sum_{c\leftarrow a} \int_{\epsilon(t)}^{1-\epsilon(t)} dz \frac{\alpha_s}{2\pi} \hat{P}_{c\leftarrow a}(z)$. For the chosen splitting, we get x_{i+1} from x_i, in analogy to eq. (4.6), as the solution of

$$\frac{\int_{\epsilon(t)}^{x_{i+1}/x_i} dz \frac{\alpha_s}{2\pi} \hat{P}_{b\leftarrow a}(z)}{\int_{\epsilon(t)}^{1-\epsilon(t)} dz \frac{\alpha_s}{2\pi} \hat{P}_{b\leftarrow a}(z)} = r_x \qquad \text{with } x_i > x_{i+1} > 0 \,, \tag{4.7}$$

where r_x is again a uniformly distributed random number in $[0, 1]$. Hence, the probability of choosing the next point (x_{i+1}, t_{i+1}) depends on the current point (x_i, t_i), but not on the points before the ith, i.e. we are dealing with a Markov process. Eqs. (4.6) and (4.7) generate t_{i+1} and x_{i+1} with the relevant probability distributions. In general, solving

$$\frac{F(y)}{F(y_{\max})} = r_y \quad \text{with } r_y \text{ flat in } [0,1], \tag{4.8}$$

generates a random number y, which is distributed according to the probability distribution

$$g(y) = \frac{1}{N_f} f(y) \,, \text{ where} \tag{4.9}$$

$$f(y) = \frac{\partial F(y)}{\partial y} \text{ and } N_f = F(y_{\max}) \,.$$

Therefore eq. (4.7) generates the momentum fraction $z = \frac{x_{i+1}}{x_i}$ with the probability distribution $\frac{1}{N_P} \frac{\alpha_s}{2\pi} \hat{P}(z)$, which gives the desired distribution for the value of x_{i+1}. Eq. (4.6) generates t_{i+1} with the probability distribution $\frac{\Delta_a(t_i)}{\Delta_a(t_{i+1})} \sum_{b\leftarrow a} \frac{1}{t_{i+1}} \int dz \frac{\alpha_s}{2\pi} \hat{P}_{b\leftarrow a}(z)$, which properly gives the probability that the parton a evolves from t_i to t_{i+1} without resolvable branching, as is given by the ratio of Sudakovs, and then branches, as given by the rest of the equation.

t_{i+1} and x_{i+1} fix two of the four momentum entries of the radiated parton. Demanding that it is on mass-shell, fixes a third. Neglecting spin correlations, we generate the azimuthal angle of the splitting randomly in $[0, 2\pi)$, to fix the last entry of the momentum. For each leg we stop the shower if $r_t < \Delta(t_i)$, then eq. (4.6) has no solution $t_{i+1} > t_0$. The solutions with $t_{i+1} < t_0$ are not considered, since these are classified as unresolvable.

4 Monte Carlo parton shower

To generate the splitting cascade for the initial state, we use a backward evolution. That is, we start from the initial lines of the hard core-process and for each parton generate stepwise the *parent* parton. For a parton b carrying a momentum fraction $x_i = x$, the probability to evolve backwards from t_i to t_{i-1} without branching is given by the function

$$\Pi(t_{i-1}, t_i, x_i) = \frac{f_b(x, t_{i-1})\Delta_b(t_i)}{f_b(x, t_i)\Delta_b(t_{i-1})},$$ (4.10)

where f_b are the parton density functions. This expression can be derived with the help of eq. (4.3). Due to the equation above, we can generated a random number r_t, uniformly distributed in $[0, 1]$, to receive t_{i-1} as the solution of

$$\Pi(t_{i-1}, t_i, x_i) = r_t.$$ (4.11)

To get the corresponding momentum fraction x_{i-1}, we solve

$$\frac{\int_0^{x_i/x_{i-1}} \mathrm{d}z \frac{\alpha_s}{2\pi} \hat{P}(z) f_b(x_i/z, t_{i-1})}{\int_0^1 \mathrm{d}z \frac{\alpha_s}{2\pi} \hat{P}(z) f_b(x_i/z, t_{i-1})} = r_x$$ (4.12)

with a random number r_x, which is equally distributed in $[0, 1]$. The azimuthal angle is again generated equidistributed in $[0, 2\pi)$.

In the algorithm above we did not account for spin correlations and for color correlations apart from the leading $1/N_c$ terms. Moreover, we assumed a strict ordering in our evolution variable t. These simplifications allowed for the probabilistic interpretation of the splitting process in terms of a simple Markov process. Obviously, this picture breaks down, when higher correlations are considered, since then the probability of a splitting would depend on more than just the last splitting. However, the aim of the shower is not to account for all higher order terms, but just for the terms, which are strongly logarithmically enhanced. As discussed in section 3.4, the leading logarithmic terms in fact arise from emission cascades ordered in t or p_t^2. In addition α_S has to be evaluated at a scale of order p_t^2. To also properly account for the NLL terms requires a close look at the exact choice of integration boundaries, ordering, scale and scheme dependence of the involved terms. A consideration to this point is done in chapter 5 for the shower in the program SHERPA 1.2.

QCD coherence

There is another attribute of QCD radiation, we can take care of in the shower. That is QCD coherence. Due to quantum interference, the nature of an additional emission from a pair of partons, which originate from the same splitting, will depend on whether its angle of emission is smaller or larger than the opening angle of the parton pair. If the emission angle is smaller than the opening angle, we can, in the large N_c limit, regard both partons as independent emitters and apply a shower algorithm. If instead the emission angle is larger than the opening angle, the new parton will be radiated coherently from both partons and just sees their total color charge, i.e. the charge of their common mother

parton. The probability to radiate a soft gluon at such an angle is found to be zero. Since the (next-to-)leading logarithms, which the shower aims to resum, just come from collinear or soft splittings, the effect of coherence can be accounted for to NLL accuracy by ordering the shower additionally in the opening angle.

Different implementations

The example of the shower algorithm above shows a general procedure how to generate splitting cascades. However, distinct Monte Carlo parton shower programs may differ from this example and among each other in various points. Examples are the choice of evolution variable as t, $p_\perp^2$ or one of them multiplied by a function of z. The shower cascade may be ordered in the evolution variable or in the opening angle. It might or might not take local recoil effects into account. The cut-offs of the integrals might be chosen differently as well as the scheme and scale of the running coupling and parton distribution functions.

Another important point is the specification of the "hard core-process". The simplest choice would be a tree level $2 \rightarrow 2$ process. For each "event" a phase space configuration of the four legs can be chosen with a chance given by the matrix element. Then the shower can be run for this event. Restricting to these kind of processes is comparatively easy to implement, but does not properly account for hard emissions. That is, because the shower just correctly accounts for the collinear and soft splittings, but not for hard ones. For the latter, eq. (3.16), which is the basis for the paton shower, is not a good approximation to the actual cross section. Although the collinear and soft emissions happen much more frequently, additional hard emissions will significantly change the energy flow of the event. For example later on we discuss the distribution of dijet event shapes as defined in chapter 7. The inclusion of properly described hard jets improves the accuracy of the shapes of the observable's distribution.

A hard emission can be correctly accounted for by the matrix element, but we have to include more than two outgoing legs. One way to do this is to consider the NLO correction to the Born $2 \rightarrow 2$ process. Here, one correctly accounts for a single hard emission and improves the normalization to NLO accuracy.

Another way to properly account for hard emissions is the CKKW approach [22]. It accounts for hard emissions of multiple jets by taking into account the tree-level $2 \rightarrow N$ matrix elements. In both approaches, new difficulties arise. A major point is to consistently merge the matrix element with the parton shower. In the following we will focus on the SHERPA program version 1.2, which uses the Catani-Seymour dipole factorization for the shower [23, 24] and an extended version of the CKKW approach [25, 22] to match tree-level multi-jet events with the shower.

4.2 Sherpa

In this section, the Monte Carlo event generator Sherpa is introduced. Sherpa [1] is a general-purpose tool for the simulation of particle collisions at high-energy colliders. It

contains very flexible tree-level matrix-element generators for the calculation of hard scattering processes within the Standard Model and various new physics models [26, 27]. The emission of additional QCD partons from the initial and final states is described through the parton-shower model of [23]. To consistently combine multi-parton matrix elements with the parton shower, a modified version [25] of the original approach [22] by Catani, Krauss, Kuhn and Webber is employed. A simple model of multiple interactions is used to account for underlying events in hadron–hadron collisions [28]. The fragmentation of partons into primary hadrons is described by a phenomenological cluster-hadronisation model [29]. A comprehensive library for simulating tau-lepton and hadron decays is provided. Where available, form-factor models and matrix elements are used to allow the inclusion of spin correlations. Effects of virtual and real QED corrections are included by the approach of Yennie, Frautschi and Suura [30].

Of all this features, we will mostly be concerned with the parton shower, which will be discussed in the next section. The matching procedure will be outlined in section 4.2.2. The numerical study, which we will perform with SHERPA in section 8, will be restricted to partonic final-states and QCD interactions.

4.2.1 The parton shower

To systematically improve event generation by including higher order corrections, also the parton shower algorithm itself must be adjusted.

At NLO infrared divergences are encountered both in the real and virtual parts. The Kinoshita-Lee-Nauenberg theorem [9, 10] guarantees their mutual cancellation for physical observables. In an explicit calculation, especially in a numeric one, it might be technically involved to achieve this cancellation. Usually one regularizes the divergences by going to $d = 4 - 2\epsilon$ dimensions. Then, they yield poles in $1/(4 - d)$ or $1/(4 - d)^2$. Since these poles follow a universal pattern for real corrections, we can choose subtraction terms, which have the same pole structure and are analytically integrable over the phase space of the additional particle. Subtracting these terms from the real corrections and adding them to the virtual corrections, the poles there cancel and both correction terms can be safely integrated numerically in four dimensions.

Motivated by this, the shower, as implemented in SHERPA 1.2 [23], uses the Catani-Seymour dipole formalism [24, 31] and its corresponding subtractions as a starting point. The catch with the subtraction terms is that they can be constructed locally from the Born matrix element. The Catani-Seymour method interprets pairs of particles as emitting particle and spectator, which are subject to a splitting kernel creating a third particle.

The essence of the Catani-Seymour method lies in the dipole factorization formula

$$|\mathcal{M}_{m+1}|^2 = \sum_{i,j} \sum_{k \neq i,j} D_{ij,k} + \sum_{i,j} \sum_{a} D_{ij}^a + \sum_{a,i} \sum_{k \neq i} D_k^{ai} + \sum_{a,i} \sum_{b \neq a} D^{ai,b} + \dots . \qquad (4.13)$$

The individual dipole contributions D provide the correct approximation of the $(m + 1)$-parton matrix element squared in the different singular regions of phase space. The dots
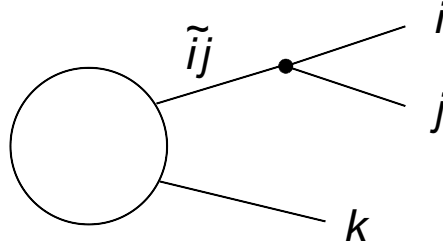
Figure 4.1: Effective diagram for the splitting of a final-state parton $\widetilde{ij}$ into partons i and j. The splitting parton is connected to a final-state spectator k through a matrix element, which is denoted by the blob.

denote finite terms, which do not have any divergence. i, j and k denote final-state partons while a and b label initial-state partons. The first sum runs over the two particles, which will be combined to the emitter by the formalism. The second sum specifies the spectators.

Of the four possible emitter-spectator combinations in eq. (4.13), which are final-final, final-initial, initial-final and initial-initial, we will focus on the combination were both are in the final state. A more complete discussion can be found in [24] for the massless case, in [31] for the massive case and in [23] for the shower implementation. In the massless case the final-final dipole contributions read

$$D_{ij,k} = \frac{1}{2p_ip_j}\langle_m 1, \ldots, \widetilde{ij}, \ldots, \tilde{k}, \ldots, m+1| \frac{\mathbf{T}_k\mathbf{T}_{ij}}{\mathbf{T}_{ij}^2}\mathbf{V}_{ij,k}|1, \ldots, \widetilde{ij}, \ldots, \tilde{k}, \ldots, m+1\rangle_m.$$

(4.14)

As shown in [23], they give the correct form of the factorized matrix element in the soft and collinear limit. In the equation above, the m-parton states are constructed from the $(m+1)$-parton matrix element by replacing the partons i and j with the emitter $\widetilde{ij}$ and the parton k with the spectator $\tilde{k}$, as illustrated in Figure 4.1. The corresponding momenta are given by

$$\tilde{p}_{ij}^\mu = p_i^\mu + p_j^\mu - \frac{y_{ij,k}}{1-y_{ij,k}}p_k^\mu \quad \text{and}$$

(4.15)

$$\tilde{p}_k^\mu = \frac{1}{1-y_{ij,k}}p_k^\mu$$

(4.16)

with the Lorentz-invariant, dimensionless quantity

$$y_{ij,k} = \frac{p_ip_j}{p_ip_j + p_ip_k + p_jp_k}.$$

(4.17)

The total four-momentum is conserved

$$p_i^\mu + p_j^\mu + p_k^\mu = \tilde{p}_{ij}^\mu + \tilde{p}_k^\mu \equiv Q$$

(4.18)

with all particles on mass shell. $\mathbf{T}$ in eq. (4.14), which are matrices in color-space, are the color charges of the emitter and spectator. $V_{ij,k}$, which are responsible for the branching

of the emitter, are matrices in its spin space. They also depend on the type of partons i and j. SHERPA does not take spin correlations into account and therefore only uses the polarization-averaged forms. For the massless case the spin averaged splitting kernels are

$$\langle V_{q_i,g_j,k}(\tilde{z}_i, y_{ij,k})\rangle = C_F \left[\frac{2}{1 - \tilde{z}_i + \tilde{z}_i y_{ij,k}} - (1 + \tilde{z}_i) \right] , \tag{4.19}$$

$$\langle V_{g_i,g_j,k}(\tilde{z}_i, y_{ij,k})\rangle = 2C_A \left[\frac{1}{1 - \tilde{z}_i + \tilde{z}_i y_{ij,k}} + \frac{1}{\tilde{z}_i + y_{ij,k} - \tilde{z}_i y_{ij,k}} - 2 + \tilde{z}_i(1 - \tilde{z}_i) \right] , \tag{4.20}$$

$$\langle V_{q_i,q_j,k}(\tilde{z}_i)\rangle = T_F \left[1 - 2\tilde{z}_i(1 - \tilde{z}_i) \right] . \tag{4.21}$$

These describe the QCD splittings $q \to qg$, $g \to gg$ and $g \to q\bar{q}$. Splittings with q interchanged with $\bar{q}$ are formally identical. In addition to $y_{ij,k}$ the splitting kernels depend on the dimensionless, Lorentz-invariant variable

$$\tilde{z}_i = 1 - \tilde{z}_j = \frac{p_i p_k}{p_i p_k + p_j p_k} . \tag{4.22}$$

In terms of the Lorentz-invariants $\tilde{z}_i$, $y_{ij,k}$ and Q^2 defined above, we can express the transverse momentum of partons i and j as

$$\vec{q}_\perp^2 = Q^2 y_{ij,k} \tilde{z}_i (1 - \tilde{z}_i) . \tag{4.23}$$

Solving this for $y_{ij,k}$, we see that for fixed $\vec{q}_\perp^2$ the Catani-Seymour splitting kernels are not singular at $\tilde{z}_i \to 0, 1$. Therefore, the soft divergence has properly been regularized. Another approximation, which is done in SHERPA, is the large N_c limit. Color correlations beyond $1/N_c$ are lost, but the color operators present in eq. (4.14) are then simply

$$-\frac{\mathbf{T}_k \mathbf{T}_{ij}}{\mathbf{T}_{ij}^2} \to 1 + \mathcal{O}\left(\frac{1}{N_c^2}\right) \quad \text{for a splitting quark,} \tag{4.24}$$

$$-\frac{\mathbf{T}_k \mathbf{T}_{ij}}{\mathbf{T}_{ij}^2} \to \frac{1}{2} + \mathcal{O}\left(\frac{1}{N_c^2}\right) \quad \text{for a splitting gluon.} \tag{4.25}$$

This can be combined to the notation

$$-\frac{\mathbf{T}_k \mathbf{T}_{ij}}{\mathbf{T}_{ij}^2} \to \frac{1}{\mathcal{N}_{ij}^{\text{spec}}} + \mathcal{O}\left(\frac{1}{N_c^2}\right) , \tag{4.26}$$

with $\mathcal{N}_{ij}^{\text{spec}}$ the number of possible spectators.

In the large N_c approximation only the planar diagrams contribute. We thus can assign a color flow to each parton configuration. Since the fixed order matrix elements are usually summed and averaged over color, this assignment is done a posteriori. However, the partons entering the shower have a well defined color and one (for a (anti)quark) or two (for a gluon) uniquely assigned color partner(s), which are the possible spectators.

With the help of the splitting kernels eqs. (4.19)-(4.21), we can express the $(m+1)$-parton matrix element as

$$|\mathcal{M}_{m+1}|^2 = |\mathcal{M}_m|^2 \sum_{ij} \sum_{k \neq ij} \frac{1}{2p_i p_j} \frac{1}{\mathcal{N}_{ij}^{\text{spec}}} 8\pi \alpha_S \langle V_{ij,k}(\tilde{z}_i, y_{ij,k}) \rangle . \tag{4.27}$$

The full phase space for the emission is given by $\tilde{z}_i \in [0,1]$, $y_{ij,k} \in [0,1]$ and $\phi \in [0, 2\pi)$. Hence, the factorized form of the $(m+1)$-parton phase space is

$$d\Phi_{m+1} = d\Phi_m \sum_{ij} \sum_{k \neq ij} \frac{2p_i p_j}{16\pi^2} \frac{dy_{ij,k}}{y_{ij,k}} d\tilde{z}_i \frac{d\phi}{2\pi} (1 - y_{ij,k}) \Theta(\tilde{z}_i(1 - \tilde{z}_i)) \Theta(y_{ij,k}(1 - y_{ij,k})) . \tag{4.28}$$

From the last two equations, we receive the factorized form of the differential cross section

$$d\sigma_{m+1} = d\sigma_m \sum_{ij} \sum_{k \neq ij} \frac{dy_{ij,k}}{y_{ij,k}} d\tilde{z}_i \frac{d\phi}{2\pi} \frac{\alpha_S}{2\pi} \frac{1}{\mathcal{N}_{ij}^{\text{spec}}} J(y_{ij,k}) \langle V_{ij,k}(\tilde{z}_i, y_{ij,k}) \rangle , \tag{4.29}$$

with the Jacobian

$$J(y_{ij,k}) = 1 - y_{ij,k} . \tag{4.30}$$

Analogously to our discussion after eq. (4.1), we can get the Sudakov form factor from the right hand side of eq. (4.29) divided by $d\sigma_m$ by integrating and exponentiating this expression. The $y_{ij,k}$ integration can be replaced by an integration over the ordering parameter, the transverse momentum, according to

$$\frac{dy_{ij,k}}{y_{ij,k}} = \frac{d\vec{q}_\perp^2}{\vec{q}_\perp^2} . \tag{4.31}$$

We cut the available phase space by the requirement of a minimal relative transverse momentum $0 < \vec{q}_{\perp,0}^2 < \vec{q}_\perp^2$ and an upper limit $\vec{q}_{\perp,\text{max}}^2 > \vec{q}_\perp^2$. While the upper limit follows directly from the kinematical constraint by the total energy, the lower limit is introduced as a cut-off to avoid the integration in the collinear region. The emissions in this region are classified as unresolvable. Due to the cuts on $\vec{q}_\perp^2$, also the $\tilde{z}_i$ integration range reduces to $\tilde{z}_i \in [z_-^o, z_+^o]$ with

$$z_{\mp}^o(\vec{q}_{\perp,\text{max}}^2, \vec{q}_{\perp,0}^2) = \frac{1}{2} \left(1 \mp \sqrt{1 - \frac{\vec{q}_{\perp,0}^2}{\vec{q}_{\perp,\text{max}}^2}} \right) . \tag{4.32}$$

For the massless final-final dipole, the Sudakov is then given by

$$\Delta_{\text{FF}}(\vec{q}_{\perp,\text{max}}^2, \vec{q}_{\perp,0}^2) = \tag{4.33}$$

$$\exp \left(\sum_{ij} \sum_{k \neq ij} \frac{1}{\mathcal{N}_{ij}^{\text{spec}}} \int_{\vec{q}_{\perp,0}^2}^{\vec{q}_{\perp,\text{max}}^2} \frac{d\vec{q}_\perp^2}{\vec{q}_\perp^2} \int_{z_-}^{z_+} d\tilde{z}_i \frac{\alpha_S(\lambda \vec{q}_\perp^2)}{2\pi} J(y_{ij,k}) \langle V_{ij,k}(\tilde{z}_i, y_{ij,k}) \rangle \right) . $$

The constant λ, which we introduced in the scale of α_S, is of order 1. We will come back to it in chapter 5.3. The Sudakov yields the non-branching probability. From it, we can generate the $\vec{q}_\perp^2$ value of the branching as in eq. (4.6) and with the splitting kernels we can generate the corresponding $\tilde{z}_i$ value as in eq. (4.7). The four-momenta of the splitting products and the spectator, which takes the recoil, are then

$$p_i = \tilde{z}_i \tilde{p}_{ij} + \frac{\vec{q}_\perp^2}{\tilde{z}_i 2\tilde{p}_{ij}\tilde{p}_k}\tilde{p}_{ij} + q_\perp \,, \tag{4.34}$$

$$p_j = (1 - \tilde{z}_i)\tilde{p}_{ij} + \frac{\vec{q}_\perp^2}{(1 - \tilde{z}_i)2\tilde{p}_{ij}\tilde{p}_k}\tilde{p}_{ij} - q_\perp \,, \tag{4.35}$$

$$p_k = (1 - y_{ij,k})\tilde{p}_k \,. \tag{4.36}$$

The corresponding expressions for other emitter-spectator combinations can be found in [23]. For them, there are some changes regarding kinematics, the definition of variables and splitting kernels. The major difference is, however, the appearance of a ratio of parton density functions in the corresponding Sudakov form factors. The corresponding expressions for the massive case can also be found in [23]. For the observables, which we want to discuss in subsequent chapters, only the massless case is relevant, because they are dominated by two jet events with large transverse momenta with respect to the masses of the light quarks.

In writing down explicit expressions for the Sudakov, a couple of scales had to be chosen. The transverse momentum $\vec{q}_\perp^2$ of the produced parton with respect to its emitter was chosen as evolution variable. Related to it the evolution cut-off $\vec{q}_{\perp,0}^2$ was introduced. To avoid the non-perturbative region, but still assign much of the phase space of particle creation to the shower, a cut-off $\vec{q}_{\perp,0}^2$ of order 1 GeV seems to be a suitable choice, as it is sufficiently separated from the Landau pole at Λ_{QCD}.

In eq. (4.33) the choice $\mu_r^2 = \lambda\vec{q}_\perp^2$ was made for the renormalization scale. An analogous choice of order $\vec{q}_\perp^2$ is made for the factorization scale μ_f, which appears as argument of the parton density functions in connection with initial sate partons.

4.2.2 The merging algorithm

The idea in SHERPA is to slice the phase space available for emissions in a hard part, which is described by the matrix element, and a collinear and soft part, which is described by the parton shower. The separation is done by a jet criterion Q, as defined in eq. (4.38). The parton shower is restricted to emissions, which obey $Q < Q_{\mathrm{cut}}$, while the matrix element is applied for $Q > Q_{\mathrm{cut}}$. An event, which is classified as a n jet event by the jet criterion, has n well separated jets of particles in the final state. Due to the matching procedure, the distribution of the n jet events is generated with the precision of the matrix element reweighted with the appropriate Sudakovs for the legs. As shown in the original paper [25], the applied merging algorithm, which is shortly summarized below, respects the logarithmic accuracy of the parton shower. For the discussion in chapter 5, we therefore

only need to look at the shower.

The generation of events in SHERPA proceeds as follows [25]:

For all jet multiplicities $n \leq N_{\max}$ and parton identities δ, the corresponding matrix element $\mathcal{M}_{n,\delta}$ and cross section $\sigma_{n,\delta}$ are calculated at tree level with the phase-space restriction $Q > Q_{\mathrm{cut}}$. The strong couplings are evaluated at a low scale to give an overestimate, which will later be reweighted. The pdfs are evaluated at the scale set by the core-process. For each matrix element the probabilities of the possible color assignments are determine in the large N_c limit, where only planar diagrams contribute.

Iteratively events are generated:

1. Select an event of jet multiplicity n with parton identities β with probability

$$P(n,\beta) = \frac{\sigma_{n,\beta}}{\sum_m^{N_{\max}} \sum_\delta \sigma_{m,\delta}} \,. \tag{4.37}$$

2. Distribute the momenta according to $|\mathcal{M}_{n,\beta}|^2$.

3. By backwards clustering determine the most probable shower history restricting to those, which correspond to a Feynman diagram, and assign the corresponding colors to the partons.

4. Accept or reject the event with a probability given by the ratio of the strong couplings in the reconstructed branchings once evaluated in the shower scheme and once in the scheme used in the matrix element.

5. Start the parton-shower evolution with suitably defined scales for intermediate and final-state particles. Intermediate partons undergo a truncated evolution. Any emission harder than Q_{cut} leads to the rejection of the event.

The last two points differ from the CKKW method, where the probability to accept or reject an event additionally depends on the Sudakovs corresponding to the reconstructed shower history and shower emissions harder than Q_{cut} are vetoed. In the case above, these Sudakovs are generated by the shower by the use of appropriate Sudakovs and the possibility to reject events. Moreover, intermediate partons undergo truncated evolution, which means that the intermediate partons are allowed to radiate between their reconstructed emissions if these emissions obey $Q < Q_{\mathrm{cut}}$. The kinematics of the following splittings is adjusted properly. This removes a short coming of the CKKW approach, where this was not possible.

The effective restriction of the shower to emissions with $Q < Q_{\mathrm{cut}}$ ensures that the n jet rates for the jet resolution Q_{cut} are generated with the rate given by the matrix element multiplied by the corresponding Sudakovs. While the procedure gives the correct rates for final-state jet multiplicities with $n \leq N_{\max}$ jets, it does not allow for more than $N_{\max}$ jets. To overcome this problem, events from a matrix element with $n = N_{\max}$ final states are treated specially, by the method of highest multiplicity treatment. After the last emission given by the matrix element has been reconstructed by the shower, the shower evolution

continues without the restriction $Q < Q_{\text{cut}}$. In this way we can account for events with higher jet multiplicities, however, the additional jets are not generated and distributed with the accuracy of the matrix element, but only with that of the shower. The choice $N_{\text{max}} = 2$ for example corresponds to the plain shower without merging. Thus, in this case not a single hard emission is generated with the accuracy of the matrix element.

The jet criterion used to distinguish the two phase space regions is for two partons i and j, whose flavor and color allows them to originate from a common mother parton $\tilde{ij}$, defined as

$$Q_{ij}^2 = 2p_i p_j \min_{k \neq i,j} \frac{2}{C_{i,j}^k + C_{j,i}^k} \,, \tag{4.38}$$

where k refers to possible spectators and for final state partons i and j we define

$$C_{i,j}^k = \frac{p_i p_k}{(p_i + p_k)p_j} - \frac{m_i^2}{2p_i p_j} \,, \text{ if } j = g \,, \tag{4.39}$$

$$C_{i,j}^k = 1 \,, \text{ else.} \tag{4.40}$$

For initial state partons a, we define

$$C_{a,j}^k = C_{(aj),j}^k \,, \tag{4.41}$$

where (aj) is created in the splitting process $a \to (aj)j$ and hence $p_{aj} = p_a - p_j$. Because Q_{ij}^2 is proportional to $p_i p_j$, it vanishes in the soft and collinear limit as needed. A more detailed discussion can be found in [25].

4.2.3 Uncertainties

There a several sources of uncertainties. Among them, we can distinguish merging-related and non-merging-related uncertainties. The latter are due to the perturbative nature of the calculations performed and would also appear without merging. They include:

- The choice of renormalization and factorization scale in the calculation of the matrix element.

- Their choice in the shower.

- The choice of parton distribution functions, which are relevant for the calculation of both the cross section and the branching probabilities for the initial-state parton shower.

- The choice of the leading-order process.

The uncertainties, which are introduced by the merging algorithm are:

- The value of the phase-space separation cut Q_{cut}.

- The choice of the jet criterion.

- The maximum number $N_{\max}$ of jets generated by the hard matrix element.

In chapter 8 we will look at the predictions from SHERPA for the distribution of dijet event shapes, which are defined in chapter 7. We will study the effect of various of these uncertainties and compare the distribution to results from resummation. Resummation is another approach to incorporate the effects of soft and collinear enhancements to all orders up to a given logarithmic accuracy. It will be introduced in chapter 6. In the next chapter, we will look at the logarithmic accuracy of the parton shower in SHERPA 1.2.

5 Logarithmic accuracy of Sherpa

The aim of this chapter is to determine the logarithmic accuracy of the parton shower as implemented in SHERPA 1.2. To this end, we will compute the anomalous Sudakov dimensions $\Gamma_{ij,k}$ in

$$\log \Delta_{\text{FF}} = -\sum_{ij} \sum_{k \neq ij} \int_{\vec{q}_{\perp,0}^2}^{\vec{q}_{\perp,\text{max}}^2} \frac{\mathrm{d}\vec{q}_\perp^2}{\vec{q}_\perp^2} \frac{\alpha_S(\vec{q}_\perp^2)}{2\pi} \Gamma_{ij,k} \,, \quad \text{i.e.} \tag{5.1}$$

$$\Gamma_{ij,k} = \int_{z_-}^{z_+} \mathrm{d}\tilde{z}_i \frac{1}{\mathcal{N}_{ij}} J(y_{ij,k}(\tilde{z}_i, \vec{q}_\perp^2, Q^2)) \langle V_{ij,k}(\tilde{z}_i, y_{ij,k}) \rangle \,. \tag{5.2}$$

Where J is given in eq. (4.30), $\langle V_{ij,k} \rangle$ in eqs. (4.19) - (4.21) and $y_{ij,k}$ in eq. (4.23). The calculation is performed up to the next-to-leading logarithmic (NLL) term. The next-to-next-to-leading logarithmic (NNLL) term is not accounted for in SHERPA, because the shower in SHERPA is strictly ordered in $\vec{q}_\perp^2$. This ordering is necessary for the implementation of the shower in terms of a Markov process. In section 3.3, we showed that emissions disordered in $\vec{p}_\perp^2$ will contribute at most a NNLL term in the expansion series.

After performing all integrations in the Sudakov, to first order in α_S, the leading logarithmic (LL) term will be proportional to $\alpha_S \log^2(\vec{q}_{\perp,0}^2/Q_{\text{hard}}^2)$, while the NLL terms will arise as $\alpha_S \log(\vec{q}_{\perp,0}^2/Q_{\text{hard}}^2)$. Where Q_{hard}^2 is some appropriate hard scale, e.g. something of order $\vec{q}_{\perp,\text{max}}^2$. If we, instead, only perform the $\tilde{z}_i$ integration, the LL term should be proportional to $\alpha_S \log(\vec{q}_\perp^2/Q_{\text{hard}}^2)$ and the NLL should be a constant independent of $\vec{q}_\perp^2$. Terms vanishing if $\vec{q}_\perp^2 \to 0$ are subleading, i.e. at most NNLL.

To simplify the notation, we write z and y instead of $\tilde{z}_i$ and $y_{ij,k}$ in the following discussion. For the z integration, there are two restrictions. On the one hand, from eq. (4.32) we have

$$z_-^o \leq z \leq z_+^o \quad \text{with}$$

$$z_\mp^o = \frac{1}{2} \left(1 \mp \sqrt{1 - \frac{\vec{q}_{\perp,0}^2}{\vec{q}_{\perp,\text{max}}^2}} \right) \,. \tag{5.3}$$

On the other hand, we also have the restriction $y \leq 1$. Using y from eq. (4.23), this leads to

$$z_-^n \leq z \leq z_+^n \quad \text{with}$$

$$z_\mp^n = \frac{1}{2} \left(1 \mp \sqrt{1 - 4\frac{\vec{q}_\perp^2}{Q^2}} \right) \,. \tag{5.4}$$

To identify, in which region which cuts is relevant, we look at the value of $\vec{q}_\perp^2$ where $z_\mp^n = z_\mp^o$. Calling it $\vec{q}_{\perp,t}^2$, we find

$$4\frac{\vec{q}_{\perp,t}^2}{Q^2} = \frac{\vec{q}_{\perp,0}^2}{\vec{q}_{\perp,\text{max}}^2} \Leftrightarrow \vec{q}_{\perp,t}^2 = \vec{q}_{\perp,0}^2 \frac{Q^2}{4\vec{q}_{\perp,m}^2} \,. \tag{5.5}$$

For $\vec{q}_\perp^2 \le \vec{q}_{\perp,t}^2$ the $z_\pm^o$ cut is more restrictive than the the $z_\pm^n$ cut. For $\vec{q}_\perp^2 \ge \vec{q}_{\perp,t}^2$ it is the other way around. To account for the existence of those two regions, we split the full $\vec{q}_\perp^2$ integral into

$$\int_{\vec{q}_{\perp,0}^2}^{\vec{q}_{\perp,t}^2} \frac{\mathrm{d}\vec{q}_\perp^2}{\vec{q}_\perp^2} \int_{z_-^o}^{z_+^o} \mathrm{d}z + \int_{\vec{q}_{\perp,t}^2}^{\vec{q}_{\perp,\text{max}}^2} \frac{\mathrm{d}\vec{q}_\perp^2}{\vec{q}_\perp^2} \int_{z_-^n}^{z_+^n} \mathrm{d}z \,. \tag{5.6}$$

If $\vec{q}_{\perp,t}^2$ is larger than $\vec{q}_{\perp,\text{max}}^2$, we only perform the first integral restricted to $[\vec{q}_{\perp,0}^2, \vec{q}_{\perp,\text{max}}^2]$. $\vec{q}_{\perp,t}^2$ will never be below $\vec{q}_{\perp,0}^2$, because the condition $y \le 1$ implies $4\vec{q}_{\perp,\text{max}}^2 \le Q^2$. Note, despite this relation, we do not expect $\vec{q}_{\perp,\text{max}}^2$ to be a function of Q^2, because Q^2 is fixed already by the parton configuration before the splitting (cf. eq. (4.18)), while in principle we want to evaluate Sudakovs with an upper bound on the $\vec{q}_\perp^2$ integration down to $\vec{q}_{\perp,0}^2$. To understand this, recall how we generate the $\vec{q}_\perp^2$ value of the splitting by finding the upper integration bound on the $\vec{q}_\perp^2$ or t integration corresponding to a random number $r \in [0,1]$ as in eq. (4.6). Thus, we see that there cannot be a fixed relation between $\vec{q}_{\perp,\text{max}}^2$ and Q^2 and both the $z_\pm^o$ and the $z_\pm^n$ cut can be relevant.

We now perform the z integrations in the two regions, identify the LL and NLL contributions and combine the results of the two regions again. The integrand in eq. (5.2) can be written as a sum of rational functions $f(z)/g(z)$ with f and g polynomials. Because neither J nor $\langle V \rangle$ do have poles in $[z_-, z_+]$, the polynomials g do not have roots there. Hence, the terms can be analytically integrated. With the shorthand notations

$$x = \frac{\vec{q}_\perp^2}{Q^2}\,, \quad s = \frac{\vec{q}_{\perp,0}^2}{\vec{q}_{\perp,\text{max}}^2}\,, \tag{5.7}$$

the results are as follows. For the splitting $q \to qg$

$$\frac{1}{\mathcal{N}_{q_{ij}}C_F} \int \mathrm{d}z J(y)\langle V_{q_i,g_j,k}(z,y)\rangle = -\frac{\log(x + (z-1)^2)}{1+x} - z - \frac{1}{2}z^2 + \tag{5.8}$$

$$-\frac{2\sqrt{x}\arctan\left(\frac{z-1}{\sqrt{x}}\right)}{1+x} - 2x\log(z-1) + \frac{(x-1)x\log(z)}{1+x} \,.$$

For the splitting $g \to gg$

$$\frac{1}{\mathcal{N}_{g_{ij}} C_A} \int \mathrm{d}z J(y) \langle V_{g_i, g_j, k}(z, y) \rangle = \tag{5.9}$$

$$\frac{1}{1+x} \left(-\frac{1}{2} \log(x + (z-1)^2) + \frac{1}{2} \log(x + z^2) - 2z + \frac{1}{2}z^2 - \frac{1}{3}z^3 \right.$$

$$- 3xz - x^2 z + \frac{1}{2}xz^2 - \frac{1}{3}xz^3 - \sqrt{x} \arctan \frac{z-1}{\sqrt{x}} - \sqrt{x} \arctan \frac{z}{\sqrt{x}}$$

$$\left. - x \log(z-1) + x \log(z) - 2x^2 \log(z-1) + 2x^2 \log(z) \right).$$

And for the splitting $g \to q\bar{q}$

$$\frac{1}{\mathcal{N}_{g_{ij}} T_F} \int \mathrm{d}z J(y) \langle V_{q_i, q_j, k}(z, y) \rangle = \tag{5.10}$$

$$z - z^2 + \frac{2}{3}z^3 + 2xz + x \log(z-1) - x \log(z).$$

The factor $\mathcal{N}_{ij}$ is 1 if the emitter is a quark and 2 if the emitter is a gluon. Its value was taken into account on the right hand side of the equations. By subtracting the right hand side of one of the last three equations evaluated at z_- from the same term evaluated at z_+, we find the corresponding anomalous Sudakov dimension Γ. Here depending on the $\vec{q}_\perp^2$ value, $z_\pm$ refers to $z_\pm^o$ or $z_\pm^n$ respectively. From all these terms, which include subleading contributions, we want to extract the LL and NLL terms. To do so, we focus on the limit of small $\vec{q}_\perp^2$ and use Taylor expansions in x or s if necessary.

5.1 Splitting quark

As first splitting, we discuss $q \to qg$, i.e. eq. (5.8). Let us start with the second integral in eq. (5.6), which uses $z_\pm^n$ from eq. (5.4) as integration boundaries. The LL contributions arises from the first term of eq. (5.8)

$$\mp \frac{\log(x + (z_\pm^n - 1)^2)}{1 + x} = \mp \frac{\log(\frac{1}{2} \mp \frac{1}{2}\sqrt{1 - 4x})}{1 + x} \approx \mp \log(\frac{1}{2} \mp \frac{1}{2} \pm x), \tag{5.11}$$

which is the LL term $-\log x$ for the upper sign and a subleading $-\log(1 - x) \approx x$ for the lower sign. The prefactor $1/(1 + x)$ is just 1 to NLL accuracy.

The NLL contribution comes from the second and third term of eq. (5.8)

$$\left[-z - \tfrac{1}{2}z^2 \right] \Big|_{z_-^n}^{z_+^n} = -\frac{3}{2} + \mathcal{O}(\text{NNLL}). \tag{5.12}$$

All other terms just give subleading contributions, as we will show now. As pointed out before, subleading terms will vanish in the limit $\vec{q}_\perp^2 \to 0$, which corresponds to $x \to 0$. To

rewrite the arctan term, we can use the relation

$$\arctan a - \arctan b = \arctan\left(\frac{a-b}{1+ab}\right), \quad \text{for } ab \geq -1.\tag{5.13}$$

In our case, we have $\frac{z_+^n - 1}{\sqrt{x}}\frac{z_-^n - 1}{\sqrt{x}} = 1 \geq -1$, thus the fourth term of eq. (5.8) gives

$$\left.\frac{-2\sqrt{x}}{1+x}\arctan\left(\frac{z-1}{\sqrt{x}}\right)\right|_{z_-^n}^{z_+^n} = \frac{-2\sqrt{x}}{1+x}\arctan\left(\frac{\sqrt{1-4x}}{\sqrt{4x}}\right).\tag{5.14}$$

Since $|\arctan y| < \pi/2$, this vanishes if $x \to 0$ and is therefore subleading. The fifth term of eq. (5.8) leads to

$$\mp 2x\log(z_\pm^n - 1) = \mp 2x\log(-z_\mp^n) = \mp i2\pi x \mp 2x\log z_\mp^n,\tag{5.15}$$

which is subleading. For the lower sign this is obvious. For the upper sign we have $-2x\log x + \mathcal{O}(\text{NNLL})$, which is subleading, because

$$\int\frac{\mathrm{d}x}{x}x\log x = \int \mathrm{d}x\log x = x(\log x - 1)\tag{5.16}$$

vanishes if x does. Thus, after the $\vec{q}_\perp^2$ integration no large logarithm arises from this term. The same reasoning holds for the sixth term of eq. (5.8).

To summarize, for the second integral in eq. (5.6), we find the anomalous Sudakov dimension

$$\Gamma_{q\to qg} = C_F\left(-\log x - \frac{3}{2}\right) + \mathcal{O}(\text{NNLL})\tag{5.17}$$

$$= C_F\left(\log\left(\frac{Q^2}{\vec{q}_\perp^2}\right) - \frac{3}{2}\right) + \mathcal{O}(\text{NNLL}).$$

The dependence on the spectator k enters only in the scale Q, which is the invariant mass of the emitter-spectator-dipole. Therefore, we did not put an explicit index k on Γ in the equation above.

To have a sensible result, we also have to look at the expression for $\Gamma_{q\to qg}$, which arises from the first integral in eq. (5.6) with the integration boundaries $z_\pm^o$. By choice of the corresponding upper $\vec{q}_\perp^2$ integration limit $\vec{q}_{\perp,t}^2$, we now have the relation $x \leq s/4$ for x and s defined in eq. (5.7). Since we expect $s/4$ to be small, we however assume $(s/4)^2 \leq x$. Then, the LL term arises from the first term in eq. (5.8) evaluated at z_+^o:

$$-\frac{\log(x + (z_+^o - 1)^2)}{1+x} = -\frac{\log(x + (z_-^o)^2)}{1+x} \approx -\log\left(x + \left(\frac{s}{4}\right)^2\right) \approx -\log x,\tag{5.18}$$

which again yields $-\log(\vec{q}_\perp^2/Q^2)$ as LL term. The corresponding term for z_-^0 gives a subleading contribution.

The NLL contribution again comes from the second and third term of eq. (5.8)

$$\left.\left(-z - \tfrac{1}{2}z^2\right)\right|_{z_-^o}^{z_+^o} = -\frac{3}{2} + \mathcal{O}(\text{NNLL}) \,. \tag{5.19}$$

All following terms in eq. (5.8) are subleading. With $\frac{z_+^o - 1}{\sqrt{x}}\frac{z_-^o - 1}{\sqrt{x}} = s/4x \geq 1 > -1$, we can use eq. (5.13) to write the fourth term as

$$\left.\frac{-2\sqrt{x}}{1+x}\arctan\left(\frac{z-1}{\sqrt{x}}\right)\right|_{z_-^o}^{z_+^o} = \frac{2\sqrt{x}}{1+x}\arctan\left(\frac{\frac{z_-^o - z_+^o}{\sqrt{x}}}{1 + \frac{s}{4x}}\right) = \frac{2\sqrt{x}}{1+x}\arctan\left(\frac{\frac{-\sqrt{1-s}}{\sqrt{x}}}{1 + \frac{s}{4x}}\right) \,. \tag{5.20}$$

This is subleading, since $|\arctan y| < \pi/2$. The contributions of the sixth term in eq. (5.8) are

$$\mp 2x\log(z_\pm^o - 1) = \mp i2\pi x \mp 2x\log z_\mp^o \,. \tag{5.21}$$

This is obviously subleading for the lower sign. For the upper sign, the leading contribution is $-2x\log s$. This is not a NLL term, but integrating it gives

$$\int_{\vec{q}_{\perp,0}^2}^{\vec{q}_{\perp,t}^2} \frac{\mathrm{d}\vec{q}_\perp^2}{\vec{q}_\perp^2} 2\frac{\vec{q}_\perp^2}{Q^2}\log\left(\frac{\vec{q}_{\perp,\text{max}}^2}{\vec{q}_{\perp,0}^2}\right) = 2\frac{\vec{q}_{\perp,t}^2 - \vec{q}_{\perp,0}^2}{Q^2}\log\left(\frac{\vec{q}_{\perp,\text{max}}^2}{\vec{q}_{\perp,0}^2}\right) \tag{5.22}$$

$$= 2\frac{\vec{q}_{\perp,0}^2}{Q^2}\left(\frac{Q^2}{4\vec{q}_{\perp,\text{max}}^2} - 1\right)\log\left(\frac{\vec{q}_{\perp,\text{max}}^2}{\vec{q}_{\perp,0}^2}\right) \,.$$

If $4\vec{q}_{\perp,\text{max}}^2$ is of order Q^2, this is negligible, because it is suppressed by a factor $\vec{q}_{\perp,0}^2/Q^2$. If $4\vec{q}_{\perp,\text{max}}^2 \ll Q^2$, i.e. $\vec{q}_{\perp,\text{max}}^2$ of order $\vec{q}_{\perp,0}^2$, its contribution is also suppressed, this time by the logarithm. Moreover, the approximation $z_-^o \approx s/4$ is rather bad in this case, because $s \approx 1$. We then instead find $z_\pm^o \approx \frac{1}{2}$ and hence the term in eq. (5.21) is subleading for all $\vec{q}_{\perp,\text{max}}^2$.

In summary, we find for the phase space region $\vec{q}_\perp^2 \in [\vec{q}_{\perp,0}^2, \vec{q}_{\perp,t}^2]$ the anomalous Sudakov dimension

$$\boxed{\Gamma_{q \to qg} = C_F\left(\log\left(\frac{Q^2}{\vec{q}_\perp^2}\right) - \frac{3}{2}\right) + \mathcal{O}(\text{NNLL}) \,.} \tag{5.23}$$

Since this is exactly the same result as eq. (5.17) for the phase space region $\vec{q}_\perp^2 \in [\vec{q}_{\perp,t}^2, \vec{q}_{\perp,\text{max}}^2]$, we can combine the sum of $\vec{q}_\perp^2$ integrals in eq. (5.6). Hence, for SHERPA 1.2, we find to NLL accuracy in the whole $\vec{q}_\perp^2$ range the anomalous Sudakov dimension for the quark as given in eq. (5.23) and (5.17). This is to NLL accuracy exactly the result, that was found in eq. (2.22) of [32]. Thus, it is correct to this accuracy.

An integration of the DGLAP splitting kernels, which are used in many other showers,

gives an analogous result. The only difference there is, that the scale Q^2 is not fixed, but is usually chosen equal to the hard scale of the core-process Q^2_{hard}. For the shower in SHERPA, instead, the scale Q^2 is fixed and arises naturally as the invariant mass of the splitting dipole. We think, that this scale choice is preferable, since all that should be relevant for the splitting, in the case of uncorrelated emissions, is the kinematic configuration of the dipole. Because the DGLAP evolution does not know about dipoles, but evolves each parton independently, there the mass of the dipole cannot arise and the scale Q^2_{hard} has to be used in place. In eq. (6.31), a similar integral as for the anomalous Sudakov dimension is solved. The scale that appears there is also the invariant mass of the dipole. This once more encourages us, to believe that the mass of the dipole is the relevant scale for Γ.

5.2 Splitting gluon

For the splitting $g \to gg$, the discussion and results will be similar to the $q \to qg$ case. This time we have to discuss the terms in eq. (5.9). The overall factor $1/(1+x)$ can be ignored to NLL accuracy.

We start with the $z^n_\pm$ case. The LL contribution comes from the first two terms of eq. (5.9)

$$\left[-\tfrac{1}{2} \log \left(x + (z-1)^2 \right) + \tfrac{1}{2} \log \left(x + z^2 \right) \right] \Bigg|_{z^n_-}^{z^n_+} = -\log \left(x + (z-1)^2 \right) \Bigg|_{z^n_-}^{z^n_+}. \tag{5.24}$$

This is exactly the term discussed in eq. (5.11). Hence, it leads to the LL contribution $-\log x$ and additional subleading terms.

The NLL contribution comes from the third to fifth term of eq. (5.9)

$$\left(-2z + \tfrac{1}{2}z - \tfrac{1}{3}z^3 \right) \Bigg|_{z_-}^{z_+} = -\frac{11}{6} + \mathcal{O}(\text{NNLL}) . \tag{5.25}$$

All other terms give subleading contributions. For the sixth to ninth term in eq. (5.9) this is obvious, because they are suppressed by a factor x. Due to the antisymmetry of the arctan, the eleventh term evaluated at z_+ and z_- gives the same contribution as the tenth term. This in turn is subleading, as shown for eq. (5.14). The twelfth and thirteenth term are subleading, which follows from the discussion of eq. (5.21). With respect to them, the last two terms in eq. (5.9) are suppressed by and additional factor x. Thus, they are subleading a fortiori.

In summary, we find for the splitting $g \to gg$ in the region $\vec{q}^2_\perp \in [\vec{q}^2_{\perp,t}, \vec{q}^2_{\perp,\text{max}}]$ the anomalous Sudakov dimension

$$\Gamma_{g \to gg} = C_A \left(\log \left(\frac{Q^2}{\vec{q}^2_\perp} \right) - \frac{11}{6} \right) + \mathcal{O}(\text{NNLL}) , \tag{5.26}$$

In the region $\vec{q}^2_\perp \in [\vec{q}^2_{\perp,0}, \vec{q}^2_{\perp,t}]$, where the $z^o_\pm$ cut is relevant, the result will be identical: The first two terms in eq. (5.9) can be combined as in eq. (5.24). According to eq. (5.18), they

then give the LL contribution $-\log(\vec{q}_\perp^2/Q^2)$. The only NLL contribution comes from the terms $-2z + z^2/2 - z^3/3$ evaluated at the boundaries $z_\pm^o$, which gives $-11/6$.

All other terms give subleading contributions. This is obvious for the sixth to ninth term in eq. (5.9), because they are suppressed by a factor x. The terms involving an arctan correspond to the ones present in the $q \to qg$ case, when we use the antisymmetry of the arctan for the second of them. Hence, they are subleading. The contributions of the twelfth and thirteenth term can be combined into $-2\frac{x}{1+x}\log(z-1)\big|_{z_-^o}^{z_+^o} + \mathcal{O}(\text{NNLL})$ and are, according to eq. (5.21) and the discussion below it, subleading. With respect to them, the two last terms in eq. (5.9) are suppressed by and additional factor x. Thus, they are subleading a fortiori.

For the phase space region $\vec{q}_\perp^2 \in [\vec{q}_{\perp,0}^2, \vec{q}_{\perp,t}^2]$ we therefore find the anomalous Sudakov dimension

$$\boxed{\Gamma_{g\to gg} = C_A\left(\log\left(\frac{Q^2}{\vec{q}_\perp^2}\right) - \frac{11}{6}\right) + \mathcal{O}(\text{NNLL}).} \tag{5.27}$$

This coincides with $\Gamma_{g\to gg}$ found for $\vec{q}_\perp^2 \in [\vec{q}_{\perp,t}^2, \vec{q}_{\perp,\text{max}}^2]$ in eq. (5.26). Thus this equation gives the anomalous Sudakov dimension for a gluon splitting to two gluons in the whole integration range. This agrees to NLL accuracy with the result found in eq. (2.23) of [32]. Thus, it is correct to this accuracy. Again, the scale Q is the invariant mass of the dipole and not an arbitrary hard scale as found for the DGLAP kernels.

For the third splitting $g \to q\bar{q}$ no LL term arises. The NLL term for both the $z_\pm^o$ and $z_\pm^n$ case comes from

$$\left(z - z^2 + \tfrac{2}{3}z^3\right)\bigg|_{z_-}^{z_+} = \frac{2}{3} + \mathcal{O}(\text{NNLL}). \tag{5.28}$$

All other terms are subleading as follows from the discussion of the other cases. Its anomalous Sudakov dimension to NLL accuracy is therefore

$$\boxed{\Gamma_{g\to q\bar{q}} = T_F\frac{1}{3} + \mathcal{O}(\text{NNLL}).} \tag{5.29}$$

Having established the correctness of the anomalous Sudakov dimensions to NLL accuracy, the same precision follows at least in the large N_c limit for the corresponding Sudakov form factor if α_S is evaluated in an appropriate scheme, up to sufficient order and at the correct scale. The appropriate choices for these will be discussed in the next section. We therefore find that the final-final Sudakov in SHERPA is correct to NLL accuracy in the large N_c limit. We restrict this statement to the large N_c limit, because SHERPA respects only leading terms in $1/N_c$ and in chapter 6 we will find that NLL terms can arise, which are not leading in $1/N_c$, and therefore negelcted in SHERPA. The finding of the relevant hard scale as the invariant mass of the splitting dipole has been discussed after eq. (5.23).

To receive a complete picture of the logarithmic accuracy of the parton shower, we would

also need to consider the Sudakovs related to dipoles, where the spectator, the emitter or both of them are in the initial state. However, in this case, the appearance of the parton density functions prevents us from performing an analogous analytical consideration. Since the kinematics and the splitting kernels are appropriately adjusted in these cases, we would, however, expect, that those Sudakovs, will also be correct to NLL accuracy in the large N_c limit if the parton density functions are evaluated at the appropriate precision. Because the matching procedure of SHERPA maintains the logarithmic accuracy of the shower, it would then follow that SHERPA produces, in the large N_c limit, events with NLL precision.

5.3 Resummed logarithms in the coupling constant

In the running coupling constant, contributions from virtual corrections are contained. When α_S is evaluated at an appropriate scale, these corrections will also lead to powers of the logarithms, which we are considering. The anomalous Sudakov dimensions Γ, which we considered in the last sections, can be written as $AL + B$. The logarithm that appears here is $L = \log \vec{q}_\perp^2/Q^2$. In the Sudakov, they appear together with $\alpha_S(\mu^2)$ inside the $\vec{q}_\perp^2$ integral. To produce the same kind of logarithms L, which appear in Γ, from the running of α_S, we see that α_S has to be evaluated at a scale proportional to $\vec{q}_\perp^2$ and expanded in terms of α_S at the scale Q^2. To do this, we apply eq. (2.34), i.e. the two-loop expansion of α_S in the $\overline{MS}$ scheme:

$$\alpha_S(\vec{q}_\perp^2) = \frac{\alpha_S(Q^2)}{1 + \frac{\beta_0}{4\pi}\alpha_S(Q^2)L} - \frac{\beta_1}{4\pi\beta_0}\frac{\alpha_S^2(Q^2)\log(1 + \frac{\beta_0}{4\pi}\alpha_S(Q^2)L)}{\left(1 + \frac{\beta_0}{4\pi}\alpha_S(Q^2)L\right)^2} + \mathcal{O}\left(\alpha_S^{k+3}L^k\right) . \quad (5.30)$$

The first summand corresponds to the one-loop expression, while the second summand gives the two loop correction. The third summand stands for higher loop corrections. Eq. (5.30) can be Taylor expanded in $\alpha_S(Q^2)L$, because the large logarithm always appears together with the strong coupling constant. The expansion will produce summands proportional to $\alpha_S^{m+n}(Q^2)L^m$. The additional factor $\alpha_S^n(Q^2)$ comes from the overall prefactors containing powers of $\alpha_S(Q^2)$. For the one-loop term $n = 1$, for the two-loop term $n = 2$ and for higher correction $n \geq 3$.

For $\alpha_S(\vec{q}_\perp^2)(AL + B)$, the one-loop expression for $\alpha_S(\vec{q}_\perp^2)$ generates summands proportional to $A\alpha_S^m(Q^2)L^m$ and $B\alpha_S^m(Q^2)L^{m-1}$. The two-loop correction, instead, leads to summands proportional to $A\alpha_S^{m+1}(Q^2)L^m$ and $B\alpha_S^{m+1}(Q^2)L^{m-1}$. The higher corrections are suppressed by additional powers of $\alpha_S(Q^2)$.

Because α_S is evaluated at the scale Q^2, the $\vec{q}_\perp^2$ integration in the Sudakov can be performed for each of these terms. As we see from eq. (3.25), a summand $\alpha_S^n(Q^2)L^m$ will become $\alpha_S^n(Q^2)\tilde{L}^{m+1}/(m+1)$, where now $\tilde{L} = \log Q^2/\vec{q}_{\perp,0}^2$ and we ignored $\log^{m+1}\vec{q}_{\perp,\max}^2/Q^2$, since it should be small compared to $\log^{m+1}\vec{q}_{\perp,0}^2/Q^2$, because typically $Q^2 \sim \vec{q}_{\perp,\max}^2$. Thus, the $\vec{q}_\perp^2$ integration produces an additional power of the large logarithm.

Therefore, the one loop expansion generates LL contributions from the $A\alpha_S(\vec{q}_\perp^2)L$ term in

the anomalous Sudakov dimension and NLL contributions from the $B\alpha_S(\vec{q}_\perp^2)$ term.[1] The two-loop expression leads to an additional NLL contribution from the $A\alpha_S(\vec{q}_\perp^2)L$ term, while from the $B\alpha_S(\vec{q}_\perp^2)$ term a NNLL contribution arises. The higher loop terms at most give NNLL corrections. We therefore see, in α_S evaluated at a scale of order $\vec{q}_\perp^2$, logarithmic corrections are resummed and to account for all LL and NLL corrections, which are leading in $1/N_c$, the two-loop expression for the running of α_S is necessary and sufficient.

NLL from correlated emissions

There is a further contribution, which we can account for, by the running of the coupling. In the parton shower, the emissions were regarded as uncorrelated apart from the ordering in $\vec{q}_\perp^2$ and kinematic dependences. However, there actually is an NLL effect of correlated emission, but as outlined in [33, 2], it can be taken into account by the expression for uncorrelated emissions if there α_S is evaluated at the scale $\vec{q}_\perp^2$ at two-loop level in the CMW scheme. The CMW scheme is related to the $\overline{MS}$ scheme by

$$\alpha_{S,\text{CMW}} = \alpha_{S,\overline{MS}} + \frac{K}{2\pi}\alpha_{S,\overline{MS}}^2\,, \tag{5.31}$$

where

$$K = \left(\frac{67}{18} - \frac{\pi^2}{6}\right)C_A - \frac{5}{9}n_f\,, \tag{5.32}$$

with n_f the number of flavors. The nice thing is that at NLL accuracy we can also obtain the CMW expression from the $\overline{MS}$ expression by changing the renormalization scale by a constant factor. Recall that in the exponent we call a the term $\alpha_S^n L^m$ a leading logarithmic term in the large logarithm L if $m = n+1$, we call it a NLL term if $m = n$, and subleading if $m < n$. Hence, if we multiply a LL term with α_S without an additional factor L, we get a NLL term. A NLL term multiplied by α_S becomes subleading. In eq. (2.34) the logarithm corresponding to the scale change $\mu_r^2 \to \lambda\mu_r^2$ is $\log\lambda$, which is not large. Hence, at NLL accuracy, we can stop the expansion of $\alpha_S(\lambda\mu_r^2)$ in terms of $\alpha_S(\mu_r^2)$ at the term $\sim \alpha_S^2$. That is, only the one-loop expression matters and from its Taylor expansion just the first two terms are relevant:

$$\alpha_S(\lambda\vec{q}_\perp^2) = \frac{\vec{q}_\perp^2}{1 + \frac{\beta_0}{4\pi}\alpha_S(\vec{q}_\perp^2)\log\lambda} = \alpha_S(\vec{q}_\perp^2)\sum_{m=0}^{\infty}\left(-\frac{\beta_0}{4\pi}\alpha_S(\vec{q}_\perp^2)\log\lambda\right)^m$$

$$= \alpha_S(\vec{q}_\perp^2) - \frac{\beta_0}{4\pi}\alpha_S^2(\vec{q}_\perp^2)\log\lambda + \mathcal{O}(\alpha_S^3)\,. \tag{5.33}$$

---

[1] As pointed out before, after all integrations in the Sudakov have been performed, summands $\alpha_S^m(Q^2)\tilde{L}^n$ are LL contributions for $n = m+1$ and NLL corrections for $m = n$, while they are subleading for $m > n$.

Demanding this to agree with the right-hand side of eq. (5.31) up to $\mathcal{O}(\alpha_S^3)$ leads to

$$-\frac{\beta_0}{4\pi} \log \lambda = \frac{K}{2\pi} \quad , \text{ i.e.} \tag{5.34}$$

$$\lambda = \exp\left(\frac{-2K}{\beta_0}\right) = \exp\left(-\frac{(67 - 3\pi^2)\frac{C_A}{3} - \frac{10 n_f}{3}}{33 - 2n_f}\right) . \tag{5.35}$$

For $SU(3)$ $C_A = 3$, thus $\lambda \approx 0.382$ for $n_f = 4$ and $\lambda \approx 0.406$ for $n_f = 5$. SHERPA adjusts its renormalization scale by this factor λ.[2] With the scale choice $\lambda \vec{q}_\perp^2$ for the running coupling in the Sudakov, we therefore account, in the large N_c limit, for all LL and NLL if the anomalous dimension is correct at this accuracy. Since we showed this for the final-final dipole in the last sections, we therefore find that the corresponding Sudakov in SHERPA is also correct to this precision. If, as we suppose, the three other dipole configurations also give the correct anomalous Sudakov dimension, the same reasoning holds for them.

---

[2]From private communication with Steffen Schumann.

6 Resummation of dijet observables

In this chapter, we will introduce the method of resummation, which takes the effect of strongly logarithmically enhanced terms into account to all orders. We will follow the considerations of [2], where the resummation method, as it is implemented in the program CAESAR, is described. As for the parton showers, a key ingredient of this method is the factorization of additional soft and collinear emissions from a hard matrix element as in eq. (3.16). In contrast to the parton showers, it, however, derives a (semi)analytical expression for the distribution of an observable by applying approximations which are justified at a given accuracy to solve the expression for the distribution explicitly. This derivation, of course, is observable specific. Here, we will look at the class of observables of dijet event shapes, which measure the distribution of momentum in the final-state of an event.

Our discussion will be structured as follows. We introduce common properties of the event shapes, which will be considered, to allow for a combined discussion of all of them. Then, we introduce the function that will be resummed. We will look at the effect of allowing additional emissions of any number of final state particles and extract their leading contributions. For these, the result for the emission of a single additional particle will be relevant. Therefore, continuing, we will consider this term. Finally, after finding the resumed result, we will discuss, how it can be combined with a fixed order calculation to improve its accuracy.

6.1 Problem specification

An event shape observable $V(k_1, k_2, ...)$ is a positive definite function of the momenta k_1, $k_2,...$ of the particles in the final state of a scattering event. It should be infrared and collinear safe. Moreover, in the limit of two narrow jets the observable should go smoothly to zero. To simplify the considerations, we will restrict to partonic final states. Then, the very limit of two narrow jets are two partons in the final state with momenta $p_1, p_2 = \{p\}$. This will be called the Born event $\mathcal{B}$. To a specific Born parton we will refer to as leg l.

For a single emission $q^{(l)}$, which is soft and collinear to leg l and has the transverse momentum $|\vec{q}_\perp^{(l)}|$, the rapidity $\eta^{(l)}$ and the azimuthal angle $\phi^{(l)}$ with respect to this leg, the observable V should be parameterizable as

$$V(\{p\}, q) = d_l \left(\frac{|\vec{q}_\perp^{(l)}|}{Q} \right)^{a_l} e^{-b_l \eta^l} g_l(\phi^{(l)}) . \tag{6.1}$$

Here Q is the hard scale of the process and $\{p\}$ are the Born momenta. a_l, b_l and $d_l g_l(\phi^{(l)})$ describe the dependence of the observable on the momentum $q^{(l)}$. The normalization

$g_l(\pi/2) = 1$ is chosen. Obviously, requiring this form, restricts the range of observables CAESAR can resum. The event shapes, which will be introduced in chapter 7, are examples of observables with these properties.

Given a procedure, which selects events with 2 or more hard jets[1], we can introduce the function $\mathcal{H}(k_1, ..., k_N)$, which is one for selected events and zero otherwise. Then we define the hard n-jet cross section as

$$\sigma = \sum_{N=2}^{\infty} \int d\Phi_N \frac{d\sigma_N}{d\Phi_N} \mathcal{H}(k_1, \ldots, k_N) , \qquad (6.2)$$

where $d\Phi_N$ is the N particle phase space and $\frac{d\sigma_N}{d\Phi_N}$ is the corresponding differential cross section.

Requiring, in addition to the selection cuts, that the value of the observable V is smaller than v, we introduce the partially integrated cross section $\Sigma(v)$ as

$$\Sigma(v) = \sum_{N=2}^{\infty} \int d\Phi_N \frac{d\sigma_N}{d\Phi_N} \Theta \left(v - V(k_1, \ldots, k_N) \right) \mathcal{H}(k_1, \ldots, k_N) . \qquad (6.3)$$

This can be rewritten in the factorized form

$$\Sigma(v) = \sum_{\delta} \int d\mathcal{B} \frac{d\sigma^{(\delta)}}{d\mathcal{B}} f_{V,\mathcal{B}}^{(\delta)}(v) \mathcal{H}(p_1, p_2) , \qquad (6.4)$$

where $\frac{d\sigma^{(\delta)}}{d\mathcal{B}}$ is the leading order differential cross section for the Born event $\mathcal{B}$ with Born momenta p_1 and p_2 in the color channel δ. We also introduced the observable dependent function $f_{V,\mathcal{B}}^{(\delta)}(v)$, whose form we will discuss in the following. For simplicity we will focus on the case of a color singlet $q\bar{q}$ pair in the final state. We will suppress the labels on $f_{V,\mathcal{B}}^{(\delta)}$. Generalizations to other color channels with more than two hard legs and to incoming partons are outlined in section 6.4.

6.2 Multiple independent emissions

Recalling what we learned about parton splitting in chapter 3, we expect a factorization of the cross section of the form

$$d\Phi_N \frac{d\sigma_N}{d\Phi_N} = d\sigma_N = d\sigma_2 \frac{1}{(N-2)!} \prod_{i=1}^{N-2} \frac{d\vec{q}_{\perp,i}^2}{\vec{q}_{\perp,i}^2} dz_i \frac{d\phi_i}{2\pi} \frac{\alpha_S}{2\pi} \hat{P}_{g \leftarrow q}(z_i) \qquad (6.5)$$

for successive emissions of real gluons from the Born legs. $\hat{P}_{g \leftarrow q}$ are the splitting kernels as defined in Table 3.1. $\vec{q}_{\perp,i}^2$, z_i and ϕ_i are respectively the transverse momentum, the

---

[1]In CAESAR this will be the k_t jet algorithm combined with a cut on the transverse momentum and rapidity.

momentum fraction and the azimuthal angle of the gluon as measured with respect to the emitting leg. The factor $1/(N-2)!$ was introduced to account for the fact that gluons are bosons. In case of $N=2$ the right hand side of eq. (6.5) is understood to be $d\sigma_2$. Because of eqs. (6.3) and (6.5), we expect f in eq. (6.4) to be given by

$$f(v) \sim \sum_{n=0}^{\infty} \frac{1}{n!} \left(\prod_{i=1}^{n} \int \frac{d\vec{q}_{\perp,i}^{2}}{\vec{q}_{\perp,i}^{2}} dz_i \frac{\alpha_S}{2\pi} \hat{P}_{g\leftarrow q}(z_i) \right) \Theta \left(v - V(\{p\}, q_1, \ldots, q_n) \right) . \qquad (6.6)$$

The equation above has to be slightly adjusted, since on the one hand it does not include virtual corrections and on the other hand the gluon can be emitted by both Born legs. We therefore take

$$\vec{q}_{\perp}^{2} = \frac{(2qp_1)(2qp_2)}{2p_1 p_2} \qquad (6.7)$$

and the momentum fractions $z^{(1)}$ and $z^{(2)}$ as given by the Sudakov decomposition of the gluon momentum q

$$q = z^{(1)} p_1 + z^{(2)} p_2 + |\vec{q}_{\perp}| \cos\phi \, n_{\text{in}} + |\vec{q}_{\perp}| \sin\phi \, n_{\text{out}} , \qquad (6.8)$$

where n_{in} and n_{out} are space-like unit vectors, which are orthogonal to p_1 and p_2 and whose vector components are in and perpendicular to the $\vec{p}_1$-$\vec{p}_2$ plane, respectively. The condition that the emission is massless implies $\vec{q}_{\perp}^{2} = z^{(1)} z^{(2)} Q_{12}^2$, where $Q_{12} = 2p_1 p_2$ is the invariant mass of the dipole. If q is approximately collinear to leg 1, i.e. if $\tan\frac{\theta_{1q}}{2} \ll \tan\frac{\theta_{12}}{2}$, where θ_{ab} is the angle between the the vector components of the momenta of a and b, then $\vec{q}_{\perp}^{2}$ and ϕ coincide with the ones defined relative to leg 1 in eq. (6.1). In this region the rapidity of the emitted gluon with respect to leg 1 is

$$\eta^{(1)} = \log\frac{2z^{(1)} E_1}{|\vec{q}_{\perp}|} = \eta + \log\frac{2E_1}{2p_1 p_2} , \qquad \eta = \frac{1}{2}\log\frac{z^{(1)}}{z^{(2)}} , \qquad (6.9)$$

where η is the rapidity of the emission in the dipole's center-of-mass frame. Analogous statements hold for emissions collinear to leg 2.

We then replace $\frac{d\vec{q}_{\perp,i}^{2}}{\vec{q}_{\perp,i}^{2}} dz_i \frac{d\phi}{2\pi} \frac{\alpha_S}{2\pi} \hat{P}_{g\leftarrow q}(z_i)$ by $[dq_i]|M(q_i)|^2$, which is a shorthand notation for

$$[dq] = \frac{d\vec{q}_{\perp}^{2}}{\vec{q}_{\perp}^{2}} dz^{(1)} \frac{d\phi}{2\pi} \qquad (6.10)$$

and

$$|M(q)|^2 = \frac{\alpha_S}{2\pi} \frac{1}{2C_F z^{(1)}} \left(z^{(1)} \hat{P}_{g\leftarrow q}(z^{(1)}) \cdot z^{(2)} \hat{P}_{g\leftarrow q}(z^{(2)}) \right) . \qquad (6.11)$$

The factor $z\hat{P}_{g\leftarrow q}/C_F$ is between one and two. It is two if z vanishes and one if z is one. The higher order corrections to the leading order (Born) cross section do not only include

real emissions, but also virtual corrections. For infrared and collinear safe observables, the infrared and collinear divergences cancel among the real and virtual corrections by means of the Kinoshita-Lee-Nauenberg theorem [9,10]. Since the virtual particles appear only in loops, their momenta are not constrained by the value v in (6.4). Thus in the soft and collinear region their contribution should correspond to the one of the real corrections as in eq. (6.6) with the adjustments as outline above and the opposite sign. For them however there is no Θ function of v. Because of this and their form, they can be written as an exponential function.

Taken together, f in eq. (6.4) is given by

$$f(v) = \exp\left(-\int [\mathrm{d}q]|M(q)|^2\right) \sum_{n=0}^{\infty} \frac{1}{n!} \left(\prod_{i=1}^{n} \int [\mathrm{d}q_i]|M(q_i)|^2\right) \times \tag{6.12}$$
$$\times \, \Theta\left(v - V(\{p\}, q_1, \ldots, q_n)\right) .$$

The first factor represents the virtual corrections. The product of integrations comes from real emissions and the sum over n and the Θ function are those, which appeared in eq. (6.3). The integrals are understood to be regularized. Moreover, we are looking at the limit of small v here, which corresponds to the soft and collinear region and allows for the factorization of the additional emissions in eq. (6.5).

Given that all emissions can be regarded as independent, eq. (6.12) holds to NLL accuracy in $\log(1/v)$. There actually is an NLL effect of correlated emission, but as outlined in [33,2], it can be taken into account by the expression for independent emissions, if α_S is evaluated at the scale $\vec{q}_\perp^2$ at two-loop-level in the CMW scheme. The latter is related to the $\overline{MS}$ scheme by eq. (5.31).

We will use the notation

$$v_i = V(\{p\}, q_i) \tag{6.13}$$

to denote the value the observable V would have if just the emission q_i and the two Born legs were present,

The emission q_i with the largest v_i we relabel as q_1. We then can write

$$\sum_{n=0}^{\infty} \frac{1}{n!} \left(\prod_{i=1}^{n} \int [\mathrm{d}q_i]|M(q_i)|^2\right) = 1 + \int [\mathrm{d}q_1]|M(q_1)|^2 \times \tag{6.14}$$
$$\times \sum_{m=0}^{\infty} \frac{1}{m!} \left(\prod_{i=2}^{m+1} \int [\mathrm{d}q_i]|M(q_i)|^2 \Theta(v_1 - v_i)\right) .$$

Let us introduce a small parameter ϵ with $0 < \epsilon \ll 1$ and $\log\frac{1}{\epsilon} \ll \log\frac{1}{v}$. We then split the sum in eq. (6.14) in two parts: the emissions satisfying $v_i \geq \epsilon v_1$ and the emissions satisfying $v_i < \epsilon v_1$. Ignoring the strongly suppressed constant term '+1' and relabeling the

q_i we can write

$$\sum_{n=0}^{\infty} \frac{1}{n!} \left(\prod_{i=1}^{n} \int [dq_i] |M(q_i)|^2 \right) = \int [dq_1] |M(q_1)|^2 \sum_{m=0}^{\infty} \frac{1}{m!} \times \tag{6.15}$$
$$\times \left(\prod_{i=2}^{m+1} \int_{\epsilon v_1}^{v_1} [dq_i] |M(q_i)|^2 \right) \sum_{n=0}^{\infty} \frac{1}{n!} \left(\prod_{j=m+2}^{n+m+1} \int_{0}^{\epsilon v_1} [dq_j] |M(q_j)|^2 \right) ,$$

with the shorthand notation for the limits

$$\int_{x}^{y} [dq_i] |M(q_i)|^2 = \int [dq_i] |M(q_i)|^2 \Theta(y - v_i) \Theta(v_i - x) . \tag{6.16}$$

This splitting is useful, because we demand the observable V to be recursively infrared and collinear (rIRC) safe. rIRC safety means that in the presence of the comparable hard emission q_1 the emissions with $v_i < \epsilon v_1$, which are much more collinear and soft than q_1, do not affect the value of the observable significantly. That is

$$V(\{p\}, q_1, \ldots, q_{m+1}, q_{m+2}, \ldots, q_{n+m+1}) = V(\{p\}, q_1, \ldots, q_{m+1}) + \mathcal{O}(\epsilon^p v_1) , \tag{6.17}$$

with some positive power p. A more explicit and extended definition is given in section 3.1 of [2]. Using this relation for V in the Θ-function of eq. (6.12) and neglecting the terms suppressed by powers of ϵ, we can write with the help of eq. (6.15)

$$f(v) = \int [dq_1] |M(q_1)|^2 \exp \left(- \int_{\epsilon v_1} [dq] |M(q)|^2 \right) \sum_{m=0}^{\infty} \frac{1}{m!} \times \tag{6.18}$$
$$\times \left(\prod_{i=2}^{m+1} \int_{\epsilon v_1}^{v_1} [dq_i] |M(q_i)|^2 \right) \Theta \left(v - V(\{p\}, q_1, \ldots, q_{m+1}) \right) .$$

Since the emissions with $v_i < \epsilon v_1$ did not appear in the common Θ-function anymore, they could be written as exponential and combined with the virtual contributions. With exception of the q_1 integral, all integrals have the lower bound ϵv_1 and are regulated by this. Moreover, the exponential term vanishes rapidly, if v_1 gets small, hence the q_1 integral will have no divergences, too. The integrand in the exponent is related to the single emission integral and we will determine its precise form to NLL accuracy in subsection 6.3. It is a function $R(\epsilon v_1)$ of the lower integral bound

$$- \int_{\epsilon v_1} [dq] |M(q)|^2 = -R(\epsilon v_1) = -R(v) - R'(v) \log \frac{v}{\epsilon v_1} + \mathcal{O}(R'') , \tag{6.19}$$

where $R'(v) = dR/d(\log 1/v)$ and we expanded R around v in a Taylor series. As long as $f(v)$ is dominated by momenta q_1 such that $v_1 \sim v$, $\log \frac{v}{\epsilon v_1}$ is not a large logarithm. Because each derivative with respect to the large logarithm $\log \frac{1}{v}$ removes a power of this logarithm, the nth correction term of the Taylor series lacks n powers of the large logarithm. Hence,

all leading logarithmic (LL) contributions are contained in $R(v)$, while NLL contributions are just contained in $R'(v) \log \frac{v}{\epsilon v_1}$ and $R(v)$. The following terms in the Taylor series are at most NNLL and can be neglected here. That $f(v)$ is in fact dominated by momenta q_1 such that $v_1 \sim v$ is discussed further in [2].

Eq. (6.19) allows to write to NLL accuracy

$$f(v) = e^{-R(v)} \mathcal{F}(v) \,, \tag{6.20}$$

where all LL are contained in the exponential factor, which is determined by the single gluon emission and the one loop virtual correction. It is a kind of Sudakov form factor, which we discussed in previous chapters. Here, however, the integration boundaries are different, since we need $V(\{p\}, q) > v$. Recalling what we learned about the interpretation of the Sudakov form factors as a non-emission probability in chapter 4, we understand the presence of this factor here. To have a value of the observable smaller than v, emissions which are such hard that they alone would already lead to $V > v$ can not be allowed. The Sudakov gives the probability that they do not occur.

In eq. (6.20), $\mathcal{F}$ is a correction factor, which accounts for the observable's dependence on multiple emissions. From eqs. (6.18) and (6.20), we get its form to NLL accuracy as

$$\mathcal{F}(v) = \int [\mathrm{d}q_1] |M(q_1)|^2 e^{-R'(v) \log \frac{v}{\epsilon v_1}} \sum_{m=0}^{\infty} \frac{1}{m!} \left(\prod_{i=2}^{m+1} \int_{\epsilon v_1}^{v_1} [\mathrm{d}q_i] |M(q_i)|^2 \right) \times \tag{6.21}$$
$$\times \, \Theta \left(v - V(\{p\}, q_1, \dots, q_{m+1}) \right) \,.$$

$\mathcal{F}$ can be evaluated in this form directly by Monte Carlo methods. To increase the efficiency and to eliminate spurious contaminations from subleading terms, a further manipulated expression is used in CAESAR. Its form and derivation is discussed in [2]. Here we will not reproduce it. Instead, we will look at the single-emission integral, because it is a central piece in eq. (6.20). The discussion in this and the following section is done for the rather simple case of a color singlet $q\bar{q}$ pair. This is the reason, why most of the complication in comparison to a parton shower only comes from including explicitly the cut v on the value of the observable V to get an analytical function of v.

6.3 Single emission integral

In this section, we will evaluate the single emission integral combined with the one loop correction. This is

$$-R(v) = \int [\mathrm{d}q] |M(q)|^2 (\Theta(v - V(\{p\}, q)) - 1) \tag{6.22}$$

$$= - \int [\mathrm{d}q] |M(q)|^2 \Theta(V(\{p\}, q) - v) \,. \tag{6.23}$$

The contribution with the Θ-function in eq. (6.22) comes from the real emission, while the -1 comes from the virtual contribution. For $V(\{p\}, q)$ we will insert eq. (6.1) assuming that

this parameterization is still sufficiently precise outside of the collinear region. Sufficiently precise means that the difference between the true observable and its parameterization just leads to a NNLL correction. We will also insert $[\mathrm{d}q]|M(q)|^2$ from eqs. (6.11) and (6.10). We will replace $\mathrm{d}z^{(1)}1/z^{(1)}$ therein by $\mathrm{d}\eta$ from eq. (6.9). Moreover, we split the η integral in two halves separated at $\eta = 0$. Emissions with $\eta = \log z^{(1)}/z^{(2)} > 0$, we assign to leg 1 and emissions with $\eta < 0$, we assign to leg 2. The full η integral, hence, becomes a sum of the terms from the two legs with η integrals limited by zero. The splitting kernel from the leg, to which the emission was not assigned, has a very small argument and we can replace the corresponding $\frac{1}{2C_F}z^{(l)}\hat{P}_{g\leftarrow q}(z^{(l)})$ by 1. With the replacements, as explained above, we rewrite $R(v)$ as

$$R(v) = \sum_{l=1}^{2} \int^{Q^2} \frac{\mathrm{d}\vec{q}_\perp^2}{\vec{q}_\perp^2} \int \eta \frac{\mathrm{d}\phi}{2\pi} \frac{\alpha_S(\vec{q}_\perp^2)}{2\pi} z^{(l)}\hat{P}_{g\leftarrow q}(z^{(l)})\times \tag{6.24}$$
$$\times \Theta(\eta)\Theta(1-z^{(l)})\Theta\left(v - d_l\left(\frac{\vec{q}_\perp^2}{Q}\right)^{a_l} e^{-b_l\eta^{(l)}} g_l(\phi)\right),$$

where Q is the hard scale of the process and for $l = 2$ we redefined $\eta = \log z^{(2)}/z^{(1)}$. The Θ-functions imply limits on the integration variables. The resulting integration region is depicted in Figure 6.1. The intersections of the limiting lines, which come from the Θ-functions in eq. (6.24) give the characteristic scales of $\vec{q}_\perp^2$. One the one hand, there are the kinematic boundaries, given by $z^{(l)} < 1$, which limit the $\vec{q}_\perp^2$ value from above. On the other hand, there is the boundary for the value of V, which limits the $\vec{q}_\perp^2$ value from below. The characteristic scales arising from the intersections of the boundaries are

$$|\vec{q}_\perp^{(l)}| \sim Q\,, \tag{6.25}$$
$$|\vec{q}_\perp^{(l)}| \sim v^{1/a_l}Q\,, \tag{6.26}$$
$$|\vec{q}_\perp^{(l)}| \sim v^{1/(a_l+b_l)}Q\,. \tag{6.27}$$

The first equation results from the limitation of $|\vec{q}_\perp^{(l)}|$ by the total energy. At the intersection of the two curves, which are due to the limitation by the value of V, η is zero and eq. (6.1) leads to the second relation. The third relation, which corresponds to the intersection of a $V = v$ curve with a $z^{(l)} < 1$ curve, also follows from eq. (6.1), but one has additionally to recall from eq. (6.9) that for $z^{(l)}$ fixed the change in $\eta^{(l)}$ is proportional to $-\log|\vec{q}_\perp^{(l)}|$. In some regions the parameterizations are just approximations to the real functions, but this does not cause trouble as long, as the difference of both just leads to a NNLL contribution, i.e. of order α_S without any large logarithm. In the region, where $z^{(1)} \sim z^{(2)} \sim 1$, eq. (6.11) is a bad approximation, but this region at most contributes $\mathcal{O}(\alpha_S)$ without logarithmic enhancements [2]. To establish numerically that the difference of the parameterized and the true form of V at most leads to a NNLL effect, CAESAR requires and tests that for collinear emissions

$$\left|\frac{\partial \log V(\{p\},q)}{\partial \log|\vec{q}_\perp^{(l)}|}\right|_{\text{fixed } z^{(l)}, \phi^{(l)}} = a_l + b_l\,, \tag{6.28}$$
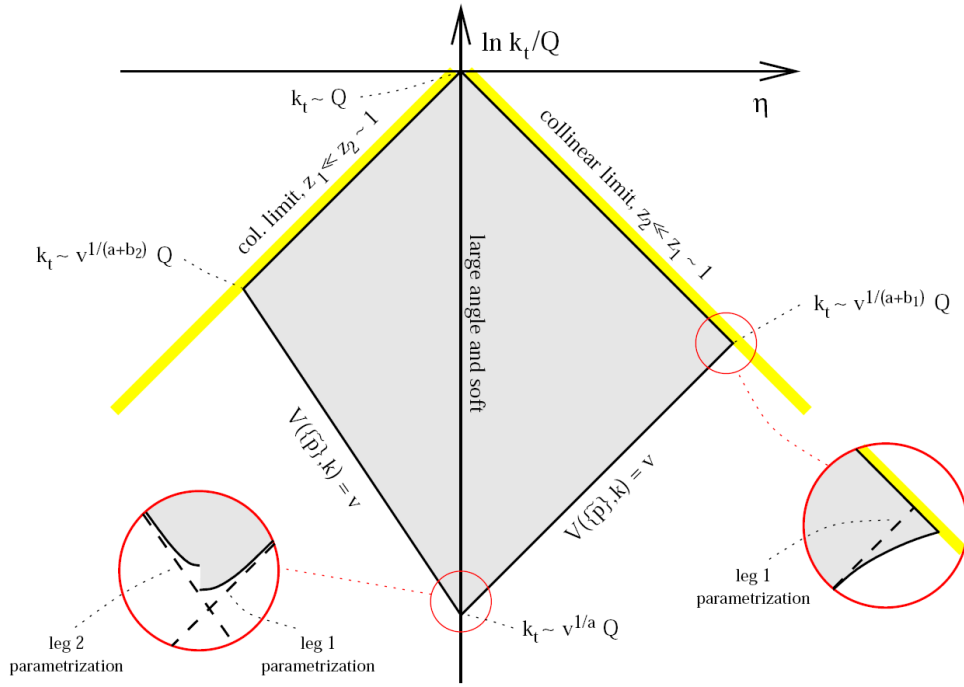
Figure 6.1: The resulting integration region for a single emission in the $\eta - \log \frac{k_t}{Q}$ plane is given by the shaded area. It is limited by the constraints on the value of the observable V and the kinematic bound on $z^{(1)}$ and $z^{(2)}$ respectively. $\eta = 0$ corresponds to the $\log \frac{k_t}{Q}$ axis, which divides the plot in the middle. Figure extracted from [2]. In our notation $k_t = |\vec{q}_\perp|$.

for soft emissions

$$
\left| \frac{\partial \log V(\{p\}, q)}{\partial \log |\vec{q}_\perp^{(l)}|} \right|_{\text{fixed } \eta^{(l)}, \, \phi^{(l)}} = a = a_1 = a_2 \, , \tag{6.29}
$$

and for both kinds of emissions V is non-zero almost everywhere. According to [2], eq. (6.29) is the condition for continuous globalness and ensures at higher orders the absence of so called non-global logarithms. These requirements are discussed further in [2]. Finally, we recall that V should vanish in the collinear and soft limit. Thus, collinear safety requires $a_l + b_l > 0$ and infrared safety $a > 0$.

We are now ready to perform the η integration in eq. (6.24). In the region, where $|\vec{q}_\perp|$ is between the values in eq. (6.26) and eq. (6.27), the upper bound for the η value comes from $\Theta\left(v - d_l \left(\frac{|\vec{q}_\perp|}{Q}\right)^{a_l} e^{-b_l \eta^{(l)}} g_l(\phi)\right)$, which limits $\eta^{(l)}$ for given $|\vec{q}_\perp|$ by $\frac{1}{b_l} \log\left[\left(\frac{|\vec{q}_\perp|}{Q}\right)^a \frac{d_l g_l(\phi)}{v}\right]$. Then, due to eq. (6.9), η is limited by $\eta^{(l)} + \log \frac{Q_{12}}{2E_l}$. Moreover, to NNLL accuracy we can approximate $z^{(l)} \hat{P}_{g \leftarrow q}(z^{(l)})$ by two, since in the largest part of this region $z^{(l)}$ is small. This will lead to the second term in eq. (6.31).

The first term of the same equation will come from the region, where $|\vec{q}_\perp|$ is limited by the values in eqs. (6.25) and (6.27). Here, the upper integration bound on η is given by $\Theta(1 - z^{(l)})$. We can change back to a $z^{(l)}$ integration and perform it. $\eta > 0$ gives the lower bound on the $z^{(l)}$ integration, e.g. for $z^{(1)}$ we have

$$
\int_{z^{(2)}}^1 \mathrm{d}z^{(1)} \hat{P}_{g \leftarrow q}(z^{(1)})/C_F = \left(\log(z^{(1)})^2 + \frac{(z^{(1)})^2 - 4z^{(1)}}{2} \right) \bigg|_{z^{(2)}}^1
$$
$$
= -\log(z^{(1)} z^{(2)}) - \frac{3}{2} + \mathcal{O}(\text{NNLL}) = \log\left(\frac{Q_{12}^2}{\vec{q}_\perp^2} \right) + B_q + \mathcal{O}(\text{NNLL}) \, . \tag{6.30}
$$

In the third step we defined $B_q = -3/2$ and used the relation $\vec{q}_\perp^2 = z^{(1)} z^{(2)} Q_{12}^2$, which we introduced after eq. (6.8). For the first term in eq. (6.31) the ϕ integration can be performed trivially, since no ϕ dependence appears there. Taken together, eq. (6.24) becomes

$$
R(v) = \sum_{l=1}^2 C_F \left[\int_{Q^2 v^{\frac{2}{a+b_l}}}^{Q^2} \frac{\mathrm{d}\vec{q}_\perp^2}{\vec{q}_\perp^2} \frac{\alpha_S(\vec{q}_\perp^2)}{2\pi} \left(\log \frac{Q_{12}^2}{\vec{q}_\perp^2} + B_q \right) + \right. \tag{6.31}
$$
$$
\left. + \int_{Q^2 v^{\frac{2}{a}}}^{Q^2 v^{\frac{2}{a+b_l}}} \frac{\mathrm{d}\vec{q}_\perp^2}{\vec{q}_\perp^2} \frac{\mathrm{d}\phi}{2\pi} \frac{\alpha_S(\vec{q}_\perp^2)}{2\pi} 2 \left(\log \frac{Q_{12}}{2E_l} + \frac{1}{b_l} \log \left[\left(\frac{|\vec{q}_\perp|}{Q} \right)^a \frac{d_l g_l(\phi)}{v} \right] \right) \right] \, .
$$

We can split the last logarithm of the second term into a ϕ independent part and into the term $\log g_l(\phi)$. Defining

$$
\log \bar{d}_l = \log d_l + \int_0^{2\pi} \frac{\mathrm{d}\phi}{2\pi} \log g_l(\phi) \tag{6.32}
$$

and

$$L \equiv \log \frac{1}{v}, \tag{6.33}$$

as well as using $\log(Q_{12}/x) = \log(Q/x) + \log(Q_{12}/Q)$, we can rewrite eq. (6.31) as

$$R(v) = \sum_{l=1}^{2} C_F \left[r_l(L) + r_l'(L) \left(\log \bar{d}_l - b_l \log \frac{2E_l}{Q} \right) + \tag{6.34} \right.$$
$$\left. - \frac{B_q}{2} T \left(\frac{L}{a + b_l} \right) \right] + 2C_F T \left(\frac{L}{a} \right) \log \frac{Q_{12}}{Q} .$$

The building blocks are the double logarithmic piece r_l, which contains all LL and some NLL contributions,

$$r_l(L) = \int_{Q^2 e^{-\frac{2L}{a+b_l}}}^{Q^2} \frac{\mathrm{d}\vec{q}_\perp^2}{\vec{q}_\perp^2} \frac{\alpha_S(\vec{q}_\perp^2)}{2\pi} \log \frac{Q^2}{\vec{q}_\perp^2} + \tag{6.35}$$
$$+ \int_{Q^2 e^{-\frac{2L}{a}}}^{Q^2 e^{-\frac{2L}{a+b_l}}} \frac{\mathrm{d}\vec{q}_\perp^2}{\vec{q}_\perp^2} \frac{\alpha_S(\vec{q}_\perp^2)}{\pi} \left(\log \frac{L}{b_l} + \log \left(\frac{|\vec{q}_\perp|}{Q} \right)^{a/b_l} \right) ,$$

its derivative $r_l' = \partial_L r_l$ and $T(L) = \int_{Q^2 e^{-2L}}^{Q^2} \frac{\mathrm{d}\vec{q}_\perp^2}{\vec{q}_\perp^2} \frac{\alpha_S(\vec{q}_\perp^2)}{\pi}$, which both do just contain single logarithmic (NLL) contributions. Taken together eqs. (6.20), (6.21) and (6.34) give the resummed function f to NLL accuracy in the large logarithm $L = \log(1/v)$.

6.4 Final form

The discussion above was for the comparatively simple case of a color singlet $q\bar{q}$ pair. However, it can be generalized to other color channels and to the case of more than 2 color charged hard legs. Then, to NLL accuracy, $R(v)$ and a reworked form of $\mathcal{F}(v)$ are adjusted by extending the sums in them over all hard, color charged legs and by adjusting the color-factors C_l and the constants B_l in the summands appropriately. Moreover, an additional NLL summand appears in $R(v)$, which accounts for the possibility that virtual corrections mix different color channels. This summand includes traces of mixing matrices and also includes terms, which are not leading in $1/N_c$.

With these modifications, the corresponding forms of eqs. (6.20), (6.21) and (6.34) give the resummed function $f_{\mathcal{B}}^{(\delta)}$ to NLL accuracy in the large logarithm $L = \log(1/v)$. With some additional work, as is done in Appendix A of [2], it is possible to eliminate spurious contamination from uncontrolled higher orders and to extract the purely LL function $Lg_1^{(\delta)}(\alpha_S L)$ and the purely NLL function $g_2^{(\delta)}(\alpha_S L)$ to write $f_{\mathcal{B}}^{(\delta)}$ in the 'standard' form

$$f_{\mathrm{NLL},\mathcal{B}}^{(\delta)}(v) = \exp \left[Lg_1^{(\delta)}(\alpha_S L) + g_{2,\mathcal{B}}^{(\delta)}(\alpha_S L) \right] . \tag{6.36}$$

From this equation it is obvious that $Lg_1^{(\delta)}(\alpha_S L)$ accounts for all terms $\alpha_S^n L^{n+1}$, which are the leading logarithms in the exponent. Its first term is the double logarithmic piece $\alpha_S L^2$. $g_{2,\mathcal{B}}^{(\delta)}(\alpha_S L)$ accounts for the terms $\alpha_S^n L^n$, which are next-to leading in the exponent. The NNLL correction in the exponent would be given by $\alpha_S g_{3,\mathcal{B}}^{(\delta)}(\alpha_S L)$ and so forth. Expanding the exponential function, we find a series $\sum_{n \geq 0} \sum_{m=n}^{2n} C_{n,m} \alpha_S^n L^m$. In this expansion the double logarithmic terms $\alpha_S^n L^{2n}$ are leading, while the terms $\alpha_S^n L^{2n-1}$ are next-to leading. NLL accuracy in the exponent implies NLL accuracy in the expansion, but not the other way round.

In eq. (6.36), $g_{2,\mathcal{B}}^{(\delta)}$ accounts for the effect of emissions from the two incoming legs i. These lead in $g_{2,\mathcal{B}}^{(\delta)}$ to the term $\sum_{i=1}^2 \log \left[p_i^{(\delta)}(x_i^{(\mathcal{B})}, v^{\frac{1}{a+b_i}} \mu_F) / p_i^{(\delta)}(x_i^{(\mathcal{B})}, \mu_F) \right]$ with $\mu_F \sim Q$. It contains the ratio of parton density functions for the longitudinal momentum fraction $x_i^{(\mathcal{B})}$ of the incoming legs i at two different factorization scales. It appears, because the pdfs at a scale μ allow for all possible incoming emissions up to $|\vec{q}_\perp| \sim \mu$, but the requirement $V < v$ for the observable, restricts collinear emissions to have $|\vec{q}_\perp| \sim v^{\frac{1}{a+b_i}} Q$. Hence, the ratio of pdfs replaces $p_i^{(\delta)}(x_i^{(\mathcal{B})}, \mu_F \sim Q)$, as used in the Born cross section, with a pdf at the correct factorization scale.

In addition to the points we just briefly mentioned above, there are some other aspects we did not discuss extensively here, but to which a more elaborated discussion can be found in [2]. Among them are the recursively infrared and collinear safety and a version of the single logarithmic function $\mathcal{F}(v)$, which is better suited for numerical integration. We also did not give the explicit form of the correction term for the mixture of different color channels.

Although we did not discuss all points extensively, our discussion highlighted important ingredients, ideas and steps of the resummation approach. We saw that contributions exponentiated and how the scale v entered the problem and appeared as large logarithm $\log(1/v)$. We also saw that the observables, which CAESAR can resum, have to fulfill a couple of rather specific requirements as those in eqs. (6.1), (6.29), (6.28) and (6.17). Hence, its region of applicability is much more limited than the one of Monte Carlo parton shower programs.

However, for multiple hard legs, the accuracy of CAESAR exceeds the precision accessible by Monte Carlo parton showers, as the latter are restricted to the large N_c limit and cannot account for the correlations of the different color channels. At this point, we also note, although the anomalous Sudakov dimensions in SHERPA, which we calculated in section 5, are correct to NLL accuracy, the Sudakov form factors have this accuracy only in the large N_c limit, as it does not provide for the correlations of different color channels.

For the simpler case of the outgoing color singlet $q\bar{q}$ pair, the accuracy of SHERPA should match the precision of CAESAR, because no mixing of color channels appears there. An interesting point to note is that the integrand of the first term in eq. (6.31) precisely corresponds to the corresponding anomalous Sudakov dimension, which we calculated for SHERPA in chapter 5. Also the hard scale is the same, i.e. the invariant mass of the dipole. Since for this term the Θ function of V was not relevant, the z integration had the full

kinematically allowed range and finding the same result here, is a consistency check to our calculation there. The second half of that equation does, of course, not appear in the Sudakovs of SHERPA, because there the cut on the value of the observable is done after an event has been generated.

As noted during our discussion, the derivation that led to eq. (6.20) is valid for small values of the observable V. This ensures that the involved radiation is not too hard. Like for the Monte Carlo parton showers programs, hard emissions are better described by a fixed order calculation. Hence, as a last point in this chapter we will to discuss, how we can match the resummed result to a fixed order calculation.

6.5 Matching to fixed order

We are aiming for a matched NLL+NLO expression for $r(v) = \frac{\Sigma(v)}{\sigma}$ with σ and $\Sigma(v)$ as defined in eqs. (6.2) and (6.3). To obtain such an expression we will follow [3]. The matched expression $r(v)$ should fulfill the following requirements: Its expansion up to relative $\mathcal{O}(\alpha_S^2)$ should reproduce the corresponding exact NLO result. In the limit of small v it should reduce to the resumed expression, which implies that in the expansion all LL and NLL logarithms are correctly accounted for. Moreover, it should respect the physical constraints $r(v_{\max}) = 1$ and $\frac{dr(v)}{dv}\big|_{v=v_{\max}} = 0$, for the maximum value $v_{\max}$ of the observable.

Various matching procedures meet these requirements. Among them is the log-R [34] and the multiplicative matching (mod-R) [35]. For both, the partially integrated cross sections $\Sigma^{(\delta)}(v)$, as defined in eq. (6.3), are calculated, once to NLO order in α_S with NLOJET++ [36] and once to NLL accuracy with CAESAR. The expressions are then combined.

While σ has the expansion

$$\sigma = \sigma_0 + \sigma_1 + \dots , \tag{6.37}$$

where σ_0 is the LO result and σ_1 is the NLO correction, $\Sigma^{(\delta)}(v)$ can be expanded in powers of α_S as

$$\Sigma^{(\delta)}(v) = \Sigma_0^{(\delta)}(v) + \Sigma_1^{(\delta)}(v) + \Sigma_2^{(\delta)}(v) + \dots , \tag{6.38}$$

for both, the fixed order result and for the resumed result. δ labels the distinct color channels. $\Sigma_0^{(\delta)}(v) = \sigma_0^{(\delta)}$, because the observable vanishes at $\mathcal{O}(\alpha_S^2)$. For the fixed order case, $\Sigma_1^{(\delta)}$ looks like a NLO term, but it contains a part $\bar{\Sigma}_1^{(\delta)}$, which is determined by the LO α_S^3 term of the differential cross section of v:[2]

$$\bar{\Sigma}_1^{(\delta)}(v) = - \int_v dv' \frac{d\Sigma_1^{(\delta)}(v)}{dv'} \tag{6.39}$$

$$\Sigma_1^{(\delta)}(v) = \sigma_1^{(\delta)} + \bar{\Sigma}_1^{(\delta)}(v) . \tag{6.40}$$

---

[2] Because the observable vanishes in case of two outgoing legs, the LO term of the differential cross section of v is determined from the tree-level diagrams with three outgoing legs. The corresponding NLO term is determined from the one-loop diagrams with three outgoing legs and from the tree-level diagrams with four outgoing legs.

A corresponding quantity $\bar{\Sigma}_2^{(\delta)}$ is determined from the NLO term of the differential cross section of v.

The resumed, partially integrated cross section will be labeled by an index r in the following. Its fixed order terms $\Sigma_{r,n}$ can be obtained by expanding eq. (6.4) in powers of α_S. To match the requirement $r(v_{\max}) = 0$, we however replace $f_{\mathcal{B}}^{(\delta)}$ therein with $\tilde{f}_{\mathcal{B}}^{(\delta)}$, which is given by eq. (6.36) with L replaced by

$$\tilde{L} \equiv \log\left(\frac{1}{x_V v} - \frac{1}{x_V v_{\max}} + 1\right) , \qquad x_V = X \cdot X_V , \qquad (6.41)$$

which leads to $\tilde{L}(v_{\max}) = 0$. Here, $\log X_V = -\frac{1}{2}\sum_{l=1}^{2}\left(\log d_l + \int \frac{\mathrm{d}\phi}{2\pi}\log g_l(\phi)\right)$, as explained in [37]. X is a parameter to test the systematic uncertainty of a matching procedure. Its default value is 1.

The log-R matching scheme defines

$$r(v) = \frac{\tilde{r}(v)}{\tilde{r}(v_{\max})} , \qquad (6.42)$$

with

$$\tilde{r}(v) = \frac{1}{\sigma_0 + \sigma_1}\left(\sum_{\delta} \tilde{\Sigma}_r^{(\delta)}(v) \exp\left[\frac{\Sigma_1^{(\delta)}(v) - \tilde{\Sigma}_{r,1}^{(\delta)}(v)}{\sigma_0^{(\delta)}}\right] \times \right. \qquad (6.43)$$

$$\left. \times \exp\left[\frac{\bar{\Sigma}_2(v) - \tilde{\Sigma}_{r,2}(v)}{\sigma_0} - \frac{1}{\sigma_0}\sum_{\delta'}\frac{(\Sigma_1^{(\delta')}(v))^2 - (\tilde{\Sigma}_{r,1}^{(\delta')}(v))^2}{2\sigma_0^{(\delta')}} + \Sigma_1^{(\text{other})}\right]\right) ,$$

where $\Sigma_1^{(\delta)}(v)$ contains the contributions of the multi-leg processes, to which a definite QCD $2 \to 2$ color channel δ can be assigned, while $\Sigma_1^{(\text{other})}(v)$ contains the contributions, for which this is not possible, e.g. a process with three final-state quarks of different flavors.

The mod-R scheme leads to

$$r(v) = \frac{1}{\sigma_0 + \sigma_1}\left(\sum_{\delta}[\tilde{\Sigma}_r^{(\delta)}(v)]^Z(\sigma_0^{(\delta)})^{1-Z}\left[1 + \frac{\Sigma_1^{(\delta)}(v) - Z\tilde{\Sigma}_{r,1}^{(\delta)}(v)}{\sigma_0^{(\delta)}} + \right.\right. \qquad (6.44)$$

$$\left.\left. + \frac{\bar{\Sigma}_2(v) - Z\tilde{\Sigma}_{r,2}(v)}{\sigma_0} - \frac{1}{\sigma_0}\sum_{\delta'}Z\tilde{\Sigma}_{r,1}^{(\delta')}(v)\frac{\Sigma_1^{(\delta')}(v) - \frac{Z+1}{2}\tilde{\Sigma}_{r,1}^{(\delta')}(v)}{\sigma_0^{(\delta')}}\right] + \Sigma_1^{(\text{other})}\right) ,$$

where $Z = \left(1 - \frac{v}{v_{\max}}\right)$. Expanding these expressions in the appropriate limits, verifies that they fulfill the requirements listed in the beginning of this section.

From the structure of the two last equations we see the general way the matching works. The resummed partially integrated cross section $\tilde{\Sigma}_r(v)$ has the expansion $\alpha_S^2 \sum_{1 \leq n \leq m} G_{nm} \alpha_S^n \tilde{L}^m$, which correctly accounts for all terms with $m \geq 2n - 1$, $n \geq 1$ and for the $n = 0$ term. The $n = 1$ term, however, does not include the term without any logarithmic

enhancements and the $n = 2$ term additionally lacks the terms, which are just enhanced by one or two logarithms. These terms are contained in the fixed order results $\Sigma_1(v)$ and $\Sigma_2(v)$ and should be added to the combined expression for $\Sigma(v)$. To extract them from the fixed order expressions $\Sigma_i(v)$ we have to subtract from the latter the terms $\Sigma_{r,i}(v)$, which are already contained in $\Sigma_r(v)$. Furthermore, at large v the considerations, which led to our expression for Σ_r, i.e. neglecting effects which are $\mathcal{O}(\text{NNLL})$ for small v, are not well motivated anymore. Thus, in this limit Σ_r should not contribute much, which is achieved by the introduction of $\tilde{L}$ as well as the variable Z in eq. (6.44).

7 Event shapes

Event shapes measure the geometrical properties of the momentum flow in an event, in especially, its deviation from pure lowest order expectations. Their distributions can, among others, be used to test QCD predictions, for new physics searches or to extract the value of the strong coupling constant. General properties and specific definitions of event shape observables in different variants will be given in this chapter. Some of them will be studied in chapter 8 numerically. In their definition, we will follow [3].

7.1 Properties

We specifically consider dijet event shapes, which measure the extend to which an event's momentum flow departs from a dijet structure. The lowest order contribution to a dijet event are two outgoing partons, which is referred to as Born event in chapter 6. In the limit of this event, the dijet event shape observables vanish smoothly. For an event with multiple hard emissions, the observables approach their maximal values, which are of order one. As meaningful observables, they have to be collinear and infrared safe. We are dealing with hadron colliders, where in an experiment the momenta longitudinal to the beam axis are unknown for both the partonic initial-state and the total final-state. Hence, the observables should be invariant under boosts longitudinal to the beam axis, which we define as the z-direction. To achieve this, the observables will be build from quantities, which are invariant under longitudinal boosts. These are the transverse momenta, squared four-vectors, rapidity differences and differences of azimuthal angles. Moreover, they will be chosen invariant with respect to rotations around the z-axis.

To match the requirements, which are needed by CAESAR for performing the resummation and which have been introduced in section 6, the observables need to have the functional form of eq. (6.1) for a single soft and collinear emission, fulfill continuous globalness as introduced with eqs. (6.28) and (6.29) and have to be recursively infrared and collinear safe as outlined for eq. (6.17). Thus, among others, the observables have to be sensible to all outgoing particles. This demand is in conflict to experimental possibilities, as one can typically just resolve particles with $|\eta| < \eta_{\max}$. For this reason, in addition to the directly global observables in section 7.2, where all particles contribute in a similar way, we will introduce other variants in section 7.3, where particles with large rapidities enter in a suppressed or indirect way.

We will use the following notations and conventions: The beam axis is chosen as z-axis. For particle i, q_i labels its four-momentum, $\vec{q}_{\perp,i} = (q_{x,i}, q_{y,i})$ its transverse components, $q_{\perp,i}$ the corresponding modulus, $\eta_i = \frac{1}{2} \ln \frac{E_i + q_{z,i}}{E_i - q_{z,i}}$ its rapidity and ϕ_i its azimuthal angle. To

reduce uncertainties from the experimental jet-energy scale and to receive dimensionless observables, the event shape variables are normalized to the hard transverse scale $Q_{\perp,X} = \sum_{i \in X} q_{\perp,i}$ with X an appropriate phase space region.

7.2 Directly global observables

Directly global observables will be labeled with the index g. All final state particles contribute to their value in the same way, such that the globalness requirement is trivially satisfied.

The **global transverse thrust** is defined as

$$T_{\perp,g} \equiv \max_{\vec{n}_T} \frac{\sum_i |\vec{q}_{\perp i} \cdot \vec{n}_T|}{\sum_i q_{\perp i}} . \tag{7.1}$$

Here and for the other global observables, the sum is over all particles in the final state. The transverse thrust axis $\vec{n}_T$ is the unit vector that maximizes $T_{\perp,g}$. Obviously, $\vec{n}_T$ lies in the transverse plain. $T_{\perp,g}$ measures the transverse energy flow along the thrust axis, which at lowest order coincides with the event axis, i.e. the axis given by the transverse components of the two outgoing particles. The observable, which will be resumed is $\tau_{\perp,g} = 1 - T_{\perp,g}$. Using $\vec{n}_T$ from (7.1), we define the directly **global thrust minor** as

$$T_{m,g} \equiv \frac{\sum_i |\vec{q}_{\perp i} \times \vec{n}_T|}{\sum_i q_{\perp i}} . \tag{7.2}$$

This quantity measures the energy flow that leaves the event plane.

An observable, which looks similar, is the **transverse spherocity** defined as

$$S_{\perp,g} \equiv \frac{\pi^2}{4} \min_{\vec{n}_\perp} \left(\frac{\sum_i |\vec{q}_{\perp i} \times \vec{n}_\perp|}{\sum_i q_{\perp i}} \right)^2 . \tag{7.3}$$

The transverse vector $\vec{n}_\perp$ that minimizes this observable is the transverse spherocity axis. The transverse spherocity measures whether the event is symmetric in the transverse plane. It is maximal for circular symmetric events.

Linearizing the transverse momentum tensor

$$M^{lin} = \sum_i \frac{1}{q_{\perp i}} \begin{pmatrix} q_{xi}^2 & q_{xi}q_{yi} \\ q_{xi}q_{yi} & q_{yi}^2 \end{pmatrix} \underset{\text{basis transf.}}{\Longleftrightarrow} \tilde{M}^{lin} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} , \tag{7.4}$$

we use its two eigenvalues $\lambda_2 \leq \lambda_1$ to define the **F-parameter** as

$$F_g = \frac{\lambda_2}{\lambda_1} . \tag{7.5}$$

This event shape variable has its its maximum 1 for events of higher multiplicity symmetrically arranged in the transverse plane.

Finally, we consider the directly global three-jet resolution parameter y_{23}. For its definition we use the exclusive variant of the **k$_t$-algorithm** [38]:

1. Given an event of n final state particles/jets $i = 1, ..., n$, define for all of them the distant measure to the beam as

$$d_{i,B}^{(n)} = \vec{q}_{\perp i}^{\,2} \tag{7.6}$$

   and the distant measure to the other particles/jets j as

$$d_{i,j}^{(n)} = \min(\vec{q}_{\perp i}^{\,2}, \vec{q}_{\perp j}^{\,2}) \frac{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}{R^2}, \quad i \neq j \tag{7.7}$$

   with the rapidity η, the azimuthal angle ϕ and the jet-radius parameter R, which sets the angular range of the jet algorithm. We use $R = 0.7$.

2. Find the minimum of all $d_{i,B}^{(n)}$ and $d_{i,j}^{(n)}$:

$$d^{(n)} = \min_{ij}(d_{i,j}^{(n)}, d_{i,B}^{(n)}) \,. \tag{7.8}$$

3. a) If $d^{(n)} = d_{i,B}^{(n)}$, remove particle/jet i from final state.

    b) If $d^{(n)} = d_{i,j}^{(n)}$, combine particles/jets i and j into a single jet by summing their four-momenta.

4. Now we are left with $n - 1$ particles/jets.

    a) If more than three particles/jets are left, go back to 1. using the decreased set of particles/jets.

    b) Else, finish the algorithm by one last clustering, such that just two jets (1, 2) are left. Define

$$P_\perp = |\vec{p}_{\perp 1}| + |\vec{p}_{\perp 2}| \,. \tag{7.9}$$

Using $P_\perp$ from eq. (7.9) and $d^{(n)}$ from eq. (7.8) for all $n \geq 3$, we define the directly **global three-jet resolution parameter** as

$$y_{23} = \frac{1}{P_\perp^2} \max_{n \geq 3}(d^{(n)}) \,. \tag{7.10}$$

7.3 Non-global and indirectly global observables

Experimentally, we can not take particles into account with rapidities $|\eta| \geq \eta_{max}$, where $\eta_{max}^{LHC} \sim 5$, $\eta_{max}^{Tev} \sim 3.5$. Motivated by this, we define the central region C as the part of phase space with $|\eta| < \eta_C$ for some appropriate η_C.

Central observables

We can define central variants of the event shapes of the last section by simply restricting the sums in eqs. (7.1) and (7.2), as well as the algorithm leading to eq. (7.10), to particles in the central region. In this way, we get the **central transverse thrust** $\tau_{\perp,C} = 1 - T_{\perp,C}$ with the new central thrust axis $\vec{n}_{T,C}$, the **central thrust minor** $T_{m,C}$ and the **central three-jet resolution threshold** $y_{23,C}$.

In addition to these observables, we introduce jet-masses and broadenings. Using the central transverse thrust axis $\vec{n}_{T,C}$, the central region C can be divided into an up part U_C containing all particles in C with $\vec{n}_{T,C} \cdot \vec{q}_{\perp i} > 0$ and a down part D_C containing all particles in C with $\vec{n}_{T,C} \cdot \vec{q}_{\perp i} < 0$. We define the normalized, squared invariant masses of the two regions as

$$\rho_{X_C} \equiv \frac{\left(\sum_{j \in X_C} q_j\right)^2}{\left(\sum_{i \in C} q_{\perp i}\right)^2} \, , \quad X = U, D \, . \tag{7.11}$$

From them, we get the **central sum of masses** $\rho_{S,C}$ and the **central heavy mass** $\rho_{H,C}$ as

$$\rho_{S,C} \equiv \rho_{U_C} + \rho_{D_C} \, , \tag{7.12}$$

$$\rho_{H,C} \equiv \max\left(\rho_{U_C}, \rho_{D_C}\right) \, . \tag{7.13}$$

Furthermore, let us introduce the mean transverse energy weighted rapidities and azimuthal angles for the up and down region

$$\eta_{X_C} \equiv \frac{\sum_{i \in X_C} q_{\perp i} \eta_i}{\sum_{i \in C_X} q_{\perp i}} \, , \quad \phi_{X_C} \equiv \frac{\sum_{i \in X_C} q_{\perp i} \phi_i}{\sum_{i \in C_X} q_{\perp i}} \, , \quad X = U, D \, . \tag{7.14}$$

Under boosts along and rotations around the beam axis, all η_i (ϕ_i) transform by the *same* summand. Hence, also η_{X_C} (ϕ_{X_C}) transforms by this summand and differences with other η_i (ϕ_i) will not change under these transformations. Thus, we can define jet broadenings

$$B_{X_C} \equiv \frac{1}{2 \sum_{j \in C} q_{\perp j}} \sum_{i \in X_C} q_{\perp i} \sqrt{(\eta_i - \eta_{X_C})^2 + (\phi_i - \phi_{X_C})^2} \, , \quad X = U, D \, , \tag{7.15}$$

which are invariant under boost along and rotations around the beam axis. The corresponding observables are the **central total** and **wide jet broadenings**

$$B_{T,C} \equiv B_{U_C} + B_{D_C} \, , \tag{7.16}$$

$$B_{W,C} \equiv \max\left(B_{U_C}, B_{D_C}\right) \, . \tag{7.17}$$

The central variables are better suited for experiments. However, they are non-global, which spoils the formal NLL+NLO accuracy of the resummation with CAESAR. To get global variants, we will add terms to them, which are sensitive to particles outside the central region.

Observables with recoil terms

Because of the conservation of transverse momentum, the **recoil term**

$$R_{\perp,C} = \frac{\left|\sum_{i \in C} \vec{q}_{\perp i}\right|}{\sum_{i \in C} q_{\perp i}} \tag{7.18}$$

is sensitive to particles outside the central region. Adding appropriate powers of this term to the central observables, produces global, **recoil enhanced variants**

$$\tau_{\perp,R} = \tau_{\perp,C} + R_{\perp,C}\,, \tag{7.19}$$

$$T_{m,R} = T_{m,C} + R_{\perp,C}\,, \tag{7.20}$$

$$y_{23,R} = y_{23,C} + R_{\perp,C}^2\,, \tag{7.21}$$

$$\rho_{S/H,R} = \rho_{S/H,C} + R_{\perp,C}\,, \tag{7.22}$$

$$B_{T/W,R} = B_{T/W,C} + R_{\perp,C}\,. \tag{7.23}$$

Unfortunately, CAESAR has large uncertainties for these observables. Hence, we introduce a further global variant of the event shapes.

Observables with exponentially suppressed forward terms

Here, we take the particles outside of the central region C into account explicitly, but in an exponentially suppressed way.

For a given event, we introduce the mean transverse energy weighted rapidity $\bar{\eta}$ of the central region

$$\bar{\eta} = \frac{\sum_{i \in C} \eta_i q_{\perp i}}{\sum_{j \in C} q_{\perp j}}\,. \tag{7.24}$$

From this, we define the exponentially suppressed forward terms as

$$\varepsilon_{\bar{C}} = \frac{1}{\sum_{i \in C} q_{\perp,i}} \sum_{i \notin C} q_{\perp,i} \exp(-|\eta_i - \bar{\eta}|)\,. \tag{7.25}$$

Note, this quantity is invariant under longitudinal boosts, because $\bar{\eta}$ transforms by the same summand as all η_i do. With the help of $\varepsilon_{\bar{C}}$, we obtain the **exponentially suppressed variants** of transverse thrust, thrust minor, three jet resolution, sum of masses, heavy mass, total jet broadening and wide jet broadening

$$\tau_{\perp,E} = \tau_{\perp,C} + \varepsilon_{\bar{C}}\,, \tag{7.26}$$

$$T_{m,E} = T_{m,C} + \varepsilon_{\bar{C}}\,, \tag{7.27}$$

$$y_{23,E} = y_{23,C} + \varepsilon_{\bar{C}}^2\,, \tag{7.28}$$

$$\rho_{S/H,E} = \rho_{S/H,C} + \varepsilon_{\bar{C}}\,, \tag{7.29}$$

$$B_{T/W,E} = B_{T/W,C} + \varepsilon_{\bar{C}}\,. \tag{7.30}$$

7.4 Accessibility

For some observables, we have defined up to four different variants, namely directly global, central, recoil enhanced and exponentially suppressed ones. To review the reasons for defining all those different variants, we recall here, which variants are accessible by a Monte Carlo parton shower, experimentally or by CAESAR.

With a Monte Carlo parton shower, such as SHERPA, all those variants are accessible and give stable numerical results. In contrast to the resummation, all observables can be calculated from a single run. Moreover, we can easily include any η-cut.

Experimentally, one usually is restricted to the region of phase space with $|\eta| < \eta_{\max}$. Hence, only the central and recoil enhanced variants are measurable. For the others, we would need to introduce η-cuts. Due to the finite resolution of the detector, there may be other particles we do not resolve. However those should be collinear or soft and therefore should not effect our collinear and infrared safe observables much.

To keep the formal NLO+NLL accuracy without a leading N_C limit, CAESAR is restricted to global observables, i.e. directly global, recoil enhanced or exponentially suppressed variants. Unfortunately, the recoil enhanced variants are numerically unstable in CAESAR. The two other global variables can not be measured in an experiment without a rapidity cut, but introducing such a cut, spoils the formal logarithmic accuracy. However, [39] suggests, that its effect should be negligible, specifically if the value of the event shape v is not too small [37], i.e.

$$v \geq v_{min} \quad \text{with} \tag{7.31}$$

$$v_{min} \sim \exp[-(a + b_{1,2})\eta_{max}] \,, \tag{7.32}$$

where a and $b_{1,2}$ are observable specific parameters defined in eq. (6.1). Including these cuts in the resummation was checked numerically in [37] and [3] to be an effect well contained in the errors. Therefore, in chapter 8, we will present results for directly global and exponentially suppressed observables, which include a η-cut.

8 Numerical study

In this chapter, we will look at the distributions obtained with SHERPA and CAESAR for the event shapes defined in chapter 7. We will consider a Tevatron scenario, i.e. $p\bar{p}$ collisions at center-of-mass energy $\sqrt{s} = 1.96$ TeV. The predictions we only perform for partonic final states. Moreover, only particles with pseudorapidity $|\eta| < 3.5$ are taken into account. The events are clustered by the SISCone jet algorithm [40] with a jet radius $R = 0.7$ and a split-merge overlap threshold $f = 0.75$. The transverse momentum p_t of each jet with respect to the beam axis is determined. The rapidities y of the two hardest jets, i.e. largest p_t, are constrained to have an absolute value of at most 0.7. For the value $p_{t,1}$ of the transverse momentum of the hardest jet, a cut is performed, which is $p_{t,1} > 50$ GeV for a low p_t sample and $p_{t,1} > 200$ GeV for a large p_t sample.

The discussion will be structured in the following way. After discussing the general shape of the observables, we will look at the systematic uncertainties of the merging procedure in SHERPA. Next, we will discuss the expected size of neglected higher order corrections in SHERPA. Finally, we will compare these results to the corresponding ones obtained by resummation with CAESAR. Since we are interested in the shape of the distributions, they will be normalized to a common scale.

As can be seen from the definition of the observables in chapter 7, they vanish in the limit of a Born event, i.e. two back-to-back particles. If in addition to the two Born legs there is a single emission with small transverse momentum $\vec{q}_\perp^2$ with respect to them, the value of the observable is proportional to a power of $\vec{q}_\perp^2$, cf. eq. (6.1). For events with multiple emissions the most relevant emission is that with the largest transverse momentum. Emissions of additional hard jets increase the value of the observable significantly. The observable's maximal value is of order 1.

To discuss the general shape and the uncertainties, we will focus on the widely known example of the transverse thrust $\tau_\perp$. We will use the variant with exponentially suppressed forward term. For the other observables, the distributions look similar and the effects of systematics and expected higher order corrections correspond to the case of the thrust.

8.1 General shape

In Figure 8.1, the distribution for the transverse thrust $\tau_{\perp,E}$, in the variant with exponentially suppressed forward terms, is given for a p_{t1}-cut of 50 GeV (a) and 200 GeV (b). The distribution vanishes rapidly for large values of the observable. This basically means that there are very few events with at least three, well separated, very hard jets, since according to the definition of the observables, it are those events, which correspond to large
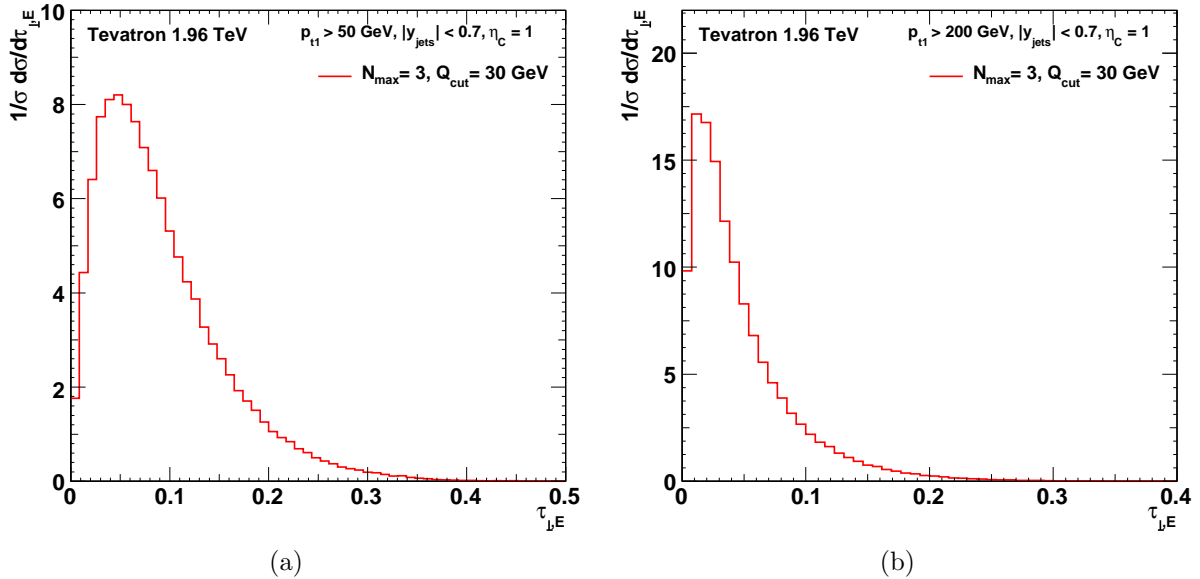
Figure 8.1: The distribution for the transverse thrust $\tau_{\perp,E}$, in the variant with exponentially suppressed forward terms. With a p_{t1}-cut of 50 GeV in (a) and of 200 GeV in (b).

values of the observables. In this region, results from fixed order matrix elements give a proper description of the distribution. Towards smaller values of the observable, but away from the Born like limit, the distribution increases - a direct and obvious consequence of the enhancement of soft and collinear radiation in QCD. Due to their enhancement, soft and collinear radiation is much more likely than hard radiation. Thus, the distribution decreases for large values of the observable.

For our discussion, the most interesting part of the distribution, however, is at very small values of the observable. Instead of increasing to arbitrary large values, as a fixed order calculation and a naive look at the splitting probability suggest, the distribution decreases for very small values of the observable. It vanishes, when the observable approaches zero. This effect is physical and in fact it is this limit that both the parton shower and the resummation aim to describe. From theory, this effect can be understood, when departing from fixed order perturbation theory and allowing for an arbitrary number of emissions, which, in fact, is done for both the parton shower and the resummation. Since additional emissions are very likely, they will happen and we never observe real Born events. A cleaner argument can be given in terms of the Sudakov form factors. The probability that a parton created at scale $\vec{p}_{\perp,m}^2$ will evolve without further emissions resolvable at scale $\vec{p}_{\perp,0}^2$ is given by the Sudakov

$$\Delta_a(\vec{p}_{\perp,m}^2) = \exp\left[-\sum_b \int_{\vec{p}_{\perp,0}^2}^{\vec{p}_{\perp,m}^2} \frac{\mathrm{d}\vec{p}_\perp^2}{\vec{p}_\perp^2} \int \mathrm{d}z \frac{\alpha_S(\vec{p}_\perp^2)}{2\pi} \hat{P}_{b\leftarrow a}(z)\right], \qquad (8.1)$$

cf. eq. (4.1) and chapter 4. For small values $\vec{p}_{\perp,0}^2$, the Sudakov vanishes faster than any

power of $\vec{p}_{\perp,0}^2$, because the leading term in the exponent is $-\alpha_S \log^2(Q^2/\vec{p}_{\perp,0}^2)$. For very small values, the observable is proportional to a power of the transverse momentum of the hardest emission. To have a value for the observable of at most v, we therefore need to include a Sudakov for both legs, where $\vec{p}_{\perp,0}^2/Q^2$ is proportional to a power of v. Hence, the Sudakov and consequently the probability to observe a small value for the observable vanishes rapidly for small values of the observable.

Because the distribution vanishes for large values v of the observable, increases for smaller v, but decreases again for very small v, it has a peak in the region of small to moderate v. The position of this peak depends on the specific observable and the selection cuts applied. Increasing the minimal value of the p_{t1}-cut shifts the peak of the distribution to smaller values of the observable as shows a comparison of Figures 8.1(a) and (b). Since we cut at rather large values of p_{t1}, the two partons entering the hard $2 \rightarrow 2$ core-process, must carry a significant amount x of the (anti-)proton's energy. Because the corresponding parton distribution functions decrease rapidly towards large x values, most of the selected events, will not have much more than the minimal energy required to survive the cut. Hence, for most events there is not enough energy left, to create additional jets, which are hard with respect to the scale set by the two hardest jets. The consequence is a shift of the distribution to smaller values of the observable, when the $\vec{p}_{\perp}^2$ cut is increased. The same argument is true if an additional hard jet comes from initial state radiation. Then we have to evaluate the parton distribution function of the emitter at a larger x scale to acquire enough energy, which allows the additional emission. Since the parton distribution functions decrease at large x, so will the chance of this to occur.

8.2 Systematic uncertainties

In this section, we will look at the systematic uncertainties of SHERPA, which are introduced through the matching procedure to multi-leg matrix elements. To this end, we will look at the dependence of the distribution of $\tau_{\perp,E}$ on $N_{\max}$, which is the maximal number of jets generated by the matrix element, and Q_{cut}, which separates the phase space regions, where the shower and the matrix element are applied, respectively. Moreover, we will check, in which way the whole distribution is build up from events with n jets in the final state.

Some of the variation between the different sets in the following discussion of systematic uncertainties is caused by statistical effects. Because most of the systematic variation is rather small, the elimination of all statistical effects would require large event samples. In addition to that, the test of all systematic uncertainties needs many different runs. Since the scale uncertainties are found in section 8.3 to be much larger than the systematic uncertainties, we decided to not eliminate all statistical effects from the systematic samples, but to only reduce them sufficiently enough to allow for the relevant observations.
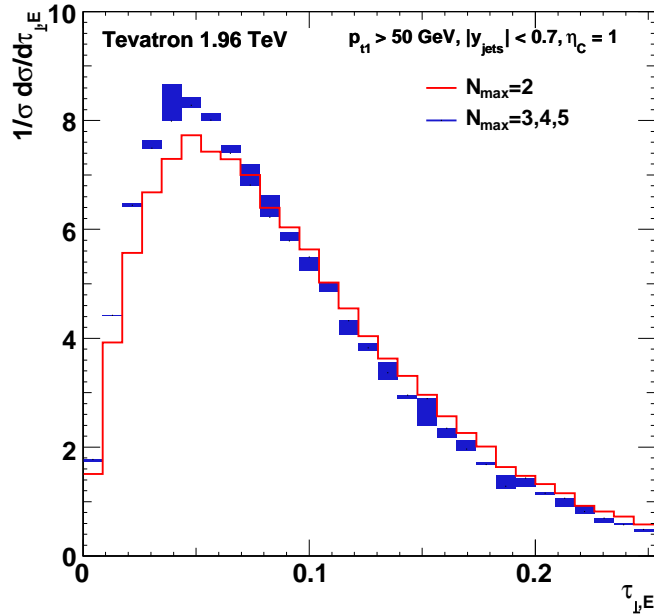
Figure 8.2: The distribution of $\tau_{\perp,E}$ for a $p_{t1} > 50$ GeV. The blue rectangulars depict the variation between the three sets with $N_{max} = 3, 4, 5$, while the red line shows the result for $N_{max} = 2$. For both $Q_{cut} = 30$ GeV was chosen.

Dependence on N_{max}

In Figure 8.2, the variation caused by the choices of $N_{max} = 2, 3, 4$ and 5 is shown for $Q_{cut} = 30$ GeV, where N_{max} is the maximal number of jets generated by the matrix element. While the distribution for $N_{max} = 2$ deviates considerably from the others, the results for $N_{max} > 2$ agree well.

The $N_{max} = 2$ case corresponds to the plain shower without merging: The matrix elements are solely used to describe the $2 \to 2$ core-process, while all emissions from its legs are generated by the shower. Due to the highest multiplicity treatment, also emissions with $Q > (>)Q_{cut}$ will be generated by the shower and the evolution will, in fact, be independent of Q_{cut} in this case. However, these hard emissions are not properly described by the shower, but would be by a matrix element. Hence, the observed deviation of the $N_{max} = 2$ sample is expected, and shows the utility of the merging approach. For improved samples we should therefore choose $N_{max} > 2$. For those samples, we observe a nice agreement in Figure 8.2 and also in corresponding plots of other observables and Q_{cut} values. We therefore find no sizeable systematic uncertainties on $N_{max} > 2$ for the dijet-event shapes.

Dependence on Q_{cut}

Next, we look at the dependence of the event shape's distribution on Q_{cut}, which separates the phase space regions, where the shower and the matrix element are applied, respectively. What choices for it are sensible? The shower evolution does not go down to arbitrarily
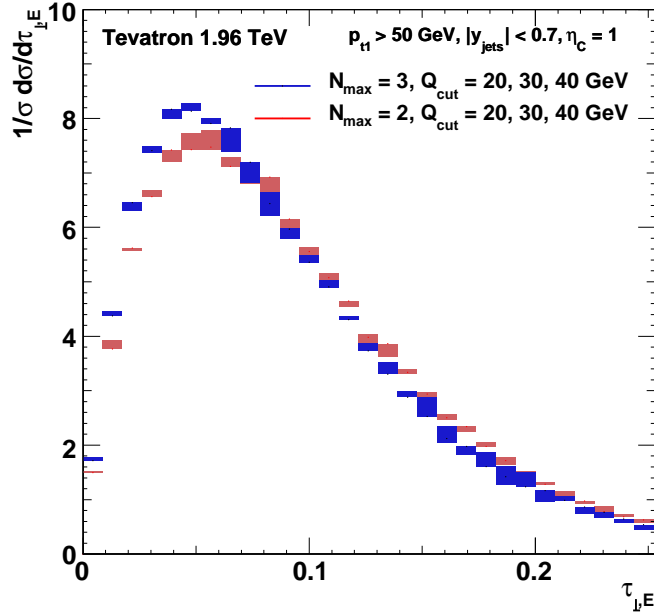
Figure 8.3: The distribution of $\tau_{\perp,E}$ for $p_{t1} > 50$ GeV. The blue rectangulars depict the variation of the $N_{\max} = 3$ sample with $Q_{\mathrm{cut}} = 20, 30, 40$ GeV. The red rectangulars the corresponding variation for the $N_{\max} = 2$ sample.

small values of Q, as we see from the cut-off on the $\vec{q}_\perp^{\,2}$ integral in eq. (4.33), which is of order 1 GeV^2. Thus Q_{cut} should be at least about 10 GeV to have some amount of phase space left, which can be filled by the shower.

In addition to that, there is also an upper bound on sensible Q_{cut} values, because we want to keep some amount of the phase space for emissions, which can be filled by the multi-leg matrix elements. The value of the p_{t1}-cut used in the event selection sets the scale, at which additional jets can still be expected with a sizable probability. So we conclude that Q_{cut} should not be much larger, than the value of the p_{t1}-cut.

Choosing Q_{cut} much larger than the value of the p_{t1}-cut would effectively reproduce the plain shower without any matching, which is also described by the choice $N_{\max} = 2$. Then of the hard emission none would be generated with the accuracy of the matrix element.

In Figure 8.3 the dependence of the distribution of $\tau_{\perp,E}$ on Q_{cut} is shown for $N_{\max} = 3$ and $p_{t1} > 50$ GeV. The variation with Q_{cut} is very small if Q_{cut} is kept in a sensible region, as outlined above. As can be seen in Figure 8.4, for too large Q_{cut} values, the distribution changes notably. This was expected, since we suppress the contribution of multi-leg matrix elements in this way and the results then effectively correspond to the results of the plain shower without matching, which is also described by the $N_{\max} = 2$ distribution. For the larger value $p_{t1,\mathrm{cut}} = 200$ GeV, these larger Q_{cut} values are still a sensible choice and the results for them agree with the lower Q_{cut} samples. Corresponding results are obtained for larger values of $N_{\max}$ and most of the other observables. For some observables, even the $N_{\max} = 2$ and the large Q_{cut} samples do not differ notably from the other samples. Those
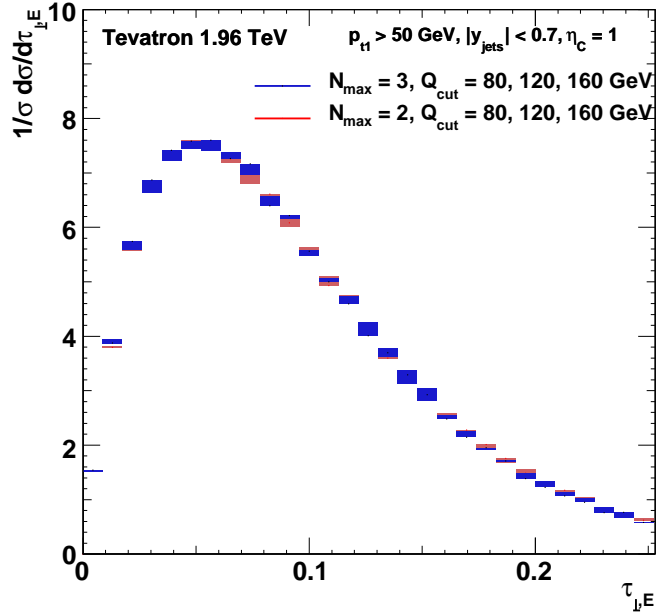
Figure 8.4: Same as Figure 8.3, but for $Q_{\mathrm{cut}} = 80, 120, 160$ GeV.

observables seem to be rather insensitive to the exact distribution of hard emissions, such that the accuracy, with which these are described by the shower, is already sufficient for those observables.

Jet decomposition

Another check, which we can do with our Monte Carlo data, is to separately compare the contributions jn, for which n jets have been generated on matrix element level. These are essentially n-jet events, as measured with the jet criterion of the matching scheme with the corresponding Q_{cut} value as resolution parameter. Due to the method of highest multiplicity treatment outlined in section 4.2.2, the sample with $n = N_{\mathrm{max}}$ however also contains events with more than n jets. Thus, the jn curves for the distributions of two different parameters $N_{\mathrm{max},1} < N_{\mathrm{max},2}$, but the same value of Q_{cut}, should not deviate much for $n < N_{\mathrm{max},1}$. The $jN_{\mathrm{max},1}$ curves, instead, will be different, because for the $N_{\mathrm{max},1}$ sample it also includes contributions from higher final-state multiplicities. We, therefore, compare the $j2$, $j3$ and $j4$ curves of the sample with $N_{\mathrm{max}} = 5$ to the corresponding jn curves from the $\tilde{N}_{\mathrm{max}} = 3, 4$ samples if $n < \tilde{N}_{\mathrm{max}}$.

This is shown in Figure 8.5 for $\tau_{\perp,E}$ with $Q_{\mathrm{cut}} = 20$ GeV. The curves agree well. If we increase the value of Q_{cut}, the contributions from the jn curves with $n > 2$ decrease, whereas the $j2$ curve increases, as shows a comparison of Figures 8.5 and 8.6. This is exactly, what is expected, since for larger values of Q_{cut}, fewer emissions will be generated by the matrix element, but the shower fills the phase space region between the new and the old Q_{cut} value. Already for $Q_{\mathrm{cut}} = 30$ GeV the jn curves with $n > 3$ only give a tiny contribution. This supports the statement, given before, that events with more than three
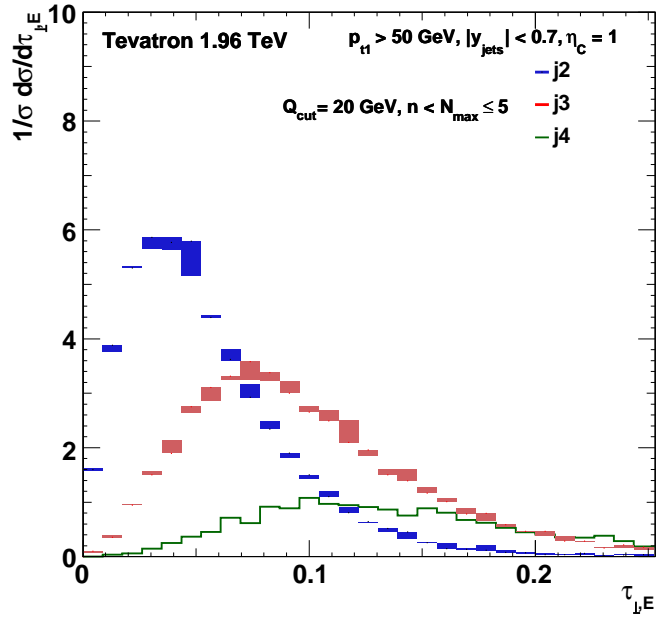
Figure 8.5: The distribution of $\tau_{\perp,E}$ for $p_{t1} > 50$ GeV. The curves depict the variation of distributions of n-jet events (jn) with $N_{\max}$. The blue rectangulars show this for $j2$ with $N_{\max} = 3, 4, 5$, the red rectangulars for $j3$ with $N_{\max} = 4, 5$ and the green curve for $j4$ with $N_{\max} = 5$. For all sets $Q_{\mathrm{cut}} = 20$ GeV was chosen.
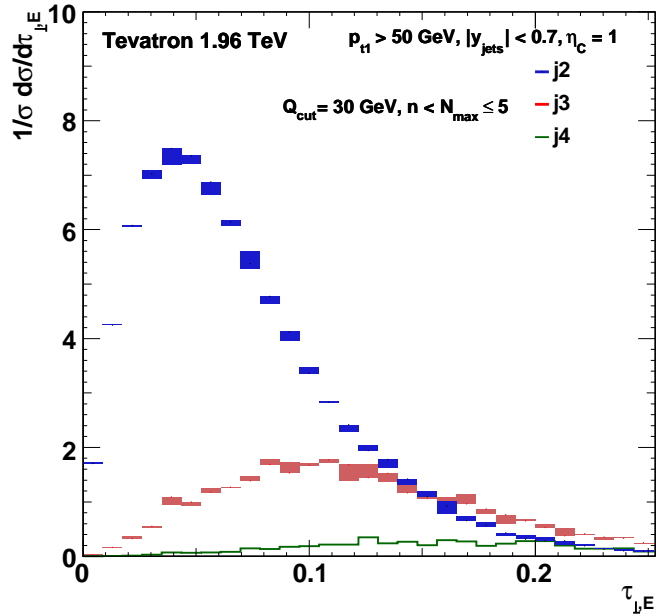

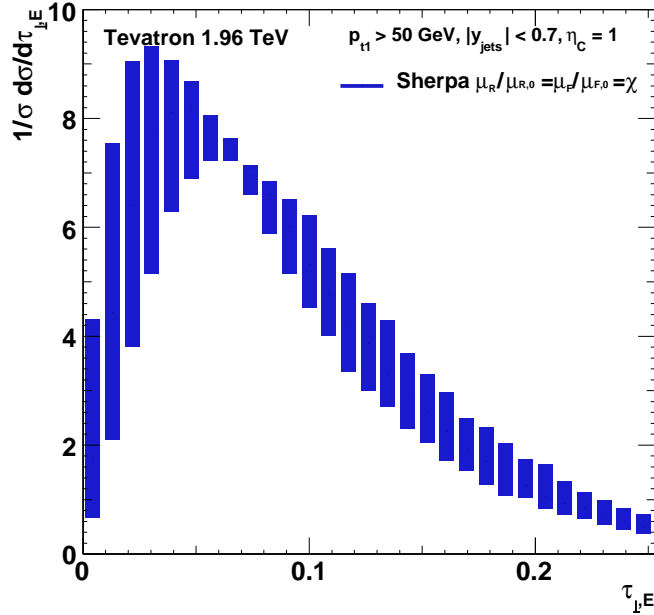Figure 8.6: Same as Figure 8.5, but with $Q_{\mathrm{cut}} = 30$ GeV.

Figure 8.7: Dependence of the $\tau_{\perp,E}$ distribution on symmetric scale variations ($\mu_r/\mu_{r,0} = \mu_f/\mu_{f,0}$) by a factor of 2 and 1/2 respectively.

hard jets effectively will never be relevant for the distribution of dijet event shapes. If we recall that the Q_{cut} variation for the full distribution in Figure 8.3 is very small, this suggests, that the phase space region, in which Q_{cut} was varied, is described similarly well by the shower and the matrix element, as required for consistency of the merging approach.

We thus find, the systematic uncertainties in SHERPA, which are introduced through the matching procedure, are small as long as we choose sensible values for N_{max} and Q_{cut}, i.e. $N_{\text{max}} > 2$ and $Q_{\text{cut}} < p_{t1,\text{cut}}$.

Since the required computing time is shortest for small N_{max} and large Q_{cut} values, we will choose $N_{\text{max}} = 3$ and $Q_{\text{cut}} = 30$ GeV in the following.

8.3 Scale uncertainties

Next, the uncertainties related to neglected higher order corrections are considered. They will be estimated by varying the renormalization scale μ_r and the factorization scale μ_f around their default values $\mu_{r,0}$ and $\mu_{f,0}$ by a factor of 2 and 1/2 respectively. The default values are for the hard $2 \rightarrow 2$ core-process are $\mu_{r,0} = \mu_{f,0} = |\vec{p}_t| = (|\vec{p}_{t,1}| + |\vec{p}_{t,2}|)/2$, where $\vec{p}_{t,1/2}$ are the transverse momenta of the two outgoing partons in the hard core-process with respect to the beam axis. For additional emissions, the default value for the factorization scale is $\mu_{f,0} = |\vec{q}_\perp|$, while the default value for the renormalization scale $\mu_{r,0} = \lambda|\vec{q}_\perp|$. Here, $\vec{q}_\perp$ is the transverse momentum of an emission with respect to the emitting leg and λ has been defined in eq. (5.35).
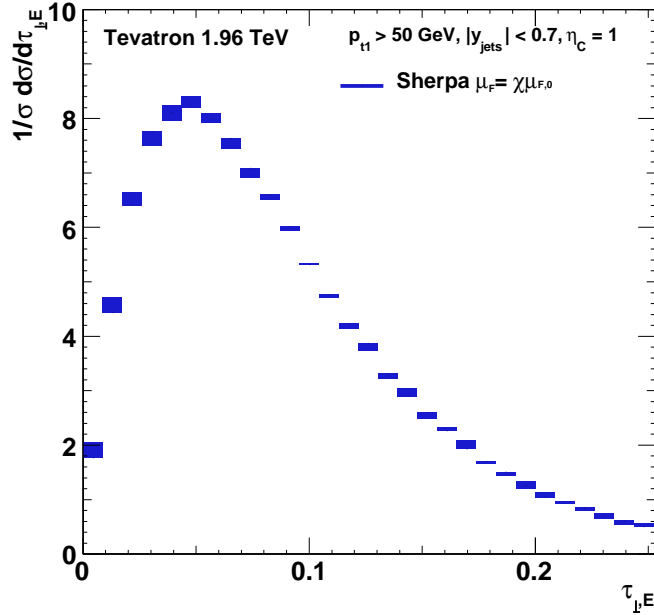
Figure 8.8: Dependence of the $\tau_{\perp,E}$ distribution on variations of the fragmentation scale μ_f by a factor of 2 and 1/2 respectively.

The dependence on scale variations is plotted for $\tau_{\perp,E}$ in Figure 8.7 for a symmetric variation, in Figure 8.8 for the variation of only μ_f and in Figure 8.9 for the variation of solely μ_r. As we observe in Figure 8.8, the shape of the distribution is essentially insensitive to μ_f variations. Of course, the total cross section changes, but the shapes are normalized to 1. In contrast, the shape of the distribution changes significantly with a change in μ_r. Increasing μ_r shifts the peak of the distribution to the left, whereas decreasing it shifts the peak to the right. Since the effect of changing μ_f is that small, the symmetric scale variations reproduce the results of the μ_r variation if both distributions are normalized to one.

Qualitatively, the effect of changing the renormalization scale by a factor of $\lambda > 0$ can easily be understood. According to the resummed one-loop expression of α_S given by the first factor of eq. (2.34), α_S will change as

$$\alpha_S(\lambda^2 \mu_r^2) = \frac{\alpha_S(\mu_r^2)}{1 + \frac{\beta_0}{4\pi}\alpha_S(\mu_r^2)\log\lambda^2} \sim \alpha_S(\mu_r^2) - \frac{\beta_0}{4\pi}\alpha_S^2(\mu_r^2)\log\lambda^2\,, \qquad (8.2)$$

where $\beta_0 > 0$ is given in eq. (2.32). This means, α_S decreases for $\lambda > 1$ and increases for $\lambda < 1$. In fixed order perturbation theory, the leading term of the cross section for a configuration with an additional particle in the final state is suppressed by a factor of α_S relatively to the cross section without this emission. Hence, decreasing α_S, i.e. increasing μ_r, reduces the chance of an emission to happen, at least if we only look at the leading terms in α_S. For $\lambda > 1$, we thus get less radiation. Since less radiation corresponds to smaller values of the event shape variable v, its distribution is shifted to the left in this case. For $\lambda < 1$, it is shifted to the right.
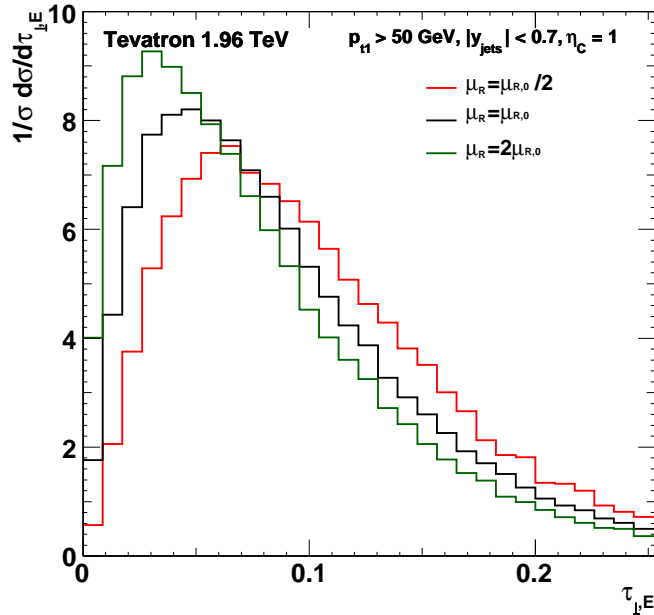
Figure 8.9: Dependence of the $\tau_{\perp,E}$ distribution on variations of the renormalization scale μ_r by a factor of 2 and 1/2 respectively. Since the normalized histograms are basically identical to Figure 8.7, we chose to not plot a band, but the single histograms here.

The same conclusion can be obtained by considering the Sudakov. The integrand in its exponent is proportional to α_S. Hence, its absolute value will increase if α_S is increased. Since the exponent is negative, the Sudakov and through it the probability to not radiate decreases in that case. This again corresponds to a shift to the right in the distribution of v.

In the following, we argue, why the variation of the factorization scale might have only a small effect. In SHERPA, the factorization scale only determines the scale, at which the parton distribution functions are evaluated, the phase space range of the $\vec{q}_\perp^2$ integration in the Sudakov, i.e. $\vec{q}_{\perp,0}^2$ and $\vec{q}_{\perp,\max}^2$ in eq. (4.33), is, however, not affected. Apart from the calculation of the tree level cross section, the parton distribution functions enter only the Sudakovs, in which the emitter, the spectator or both of them are in the initial state. There, they always appear as a *ratio* of two parton distribution functions of the same parton type evaluated at two different momentum fractions. Such that a part of the μ_f dependence will cancel between the two parton distribution functions there. In particular, because in the relevant region of moderate to large x,[1] a change of μ_f by a factor of 2 or 1/2 does not impact the general shape of the parton distribution function for a given parton type much and, hence, it is merely a parton type dependent factor, by which the parton distribution functions are changed in that region. Such a factor would cancel between the parton density functions in the Sudakov and for the core-process merely affect its total

---

[1]The $p_{t1} > 50$ GeV cut requires approximately $x > 0.05$.

value, but not its distribution. This might explain that the variation of the event shape's distributions is rather insensitive to the μ_f variation performed.

One effect, the change of factorization scale should have, apart from changing the total cross section, is to reweight the contributions of the different partonic subchannels. For a larger value of μ_f, the gluon-gluon fusion channel, for example, should become more important. However, it appears that the difference in the shapes of the subchannels is not such drastically that the variation of μ_f by a factor of 2 or 1/2 has a significant impact on the total shape.

The uncertainties related to scale variations are significantly larger than the systematic uncertainties discussed before if the improper choices $N_{\max} = 2$ and $Q_{\mathrm{cut}} \gtrsim p_{t1,\mathrm{cut}}$ are ignored. Hence, it will be sufficient to include the former ones, when we compare the results of SHERPA with the results from CAESAR in section 8.5. Moreover, since the effect of the μ_f variation is that small, it will be sufficient to consider symmetric scale variations for SHERPA.

We hope to gain a meaningful estimate for uncertainties in SHERPA in this way. We, however, did not test all possible sources of uncertainties, because some are hard to assess. For example, the effect of ignoring non-leading terms in $1/N_c$ cannot be tested in Sherpa, as a parton shower intrinsically cannot contribute for these terms. Another example is the choice of the jet criterion, which we at most tested indirectly by varying Q_{cut}.

8.4 Results from Caesar

In [3], results for global and indirectly global variants of the observables defined in chapter 7 have been obtained. The distributions were determined to NLO+NLL accuracy. For the resummation part CAESAR [2] was used, while the fixed order calculation was performed by NLOJET++ 3.0 [36]. For completeness, we will shortly review the sources and typical sizes of uncertainties in this approach. A more extended discussion can be found in the original paper. From the matching process, as outlined in section 6.5, systematic uncertainties arise from both the choice of the matching scheme and the scale X in eq. (6.41). In [3], they are estimated by comparing results from two different matching schemes, which are mod-R and log-R matching as defined in section 6.5, and by varying the scale X between 0.5 and 2. The uncertainties from neglected higher order contributions are estimated by the variation of the factorization and the renormalization scales by a factor of 2 and 1/2 around their default value $|\vec{p}_t| = (|\vec{p}_{t,1}| + |\vec{p}_{t,2}|)/2$, where $|\vec{p}_{t,1}|$ and $|\vec{p}_{t,2}|$ are the transverse momenta of the two hardest jets. Typical effects of these variations are shown for the example of $T_{m,g}$ in Figure 8.10. The general form of the shape corresponds to the one discussed for the SHERPA prediction for $\tau_{\perp,E}$ in section 8.1. In contrast to SHERPA, the systematic uncertainties are of comparable size as the ones from scale variations. Therefore, for the uncertainty band of the CAESAR sample in the next section, all combinations of separate variations have been considered. This also includes asymmetric scale variations ($\mu_r \neq \mu_f$).
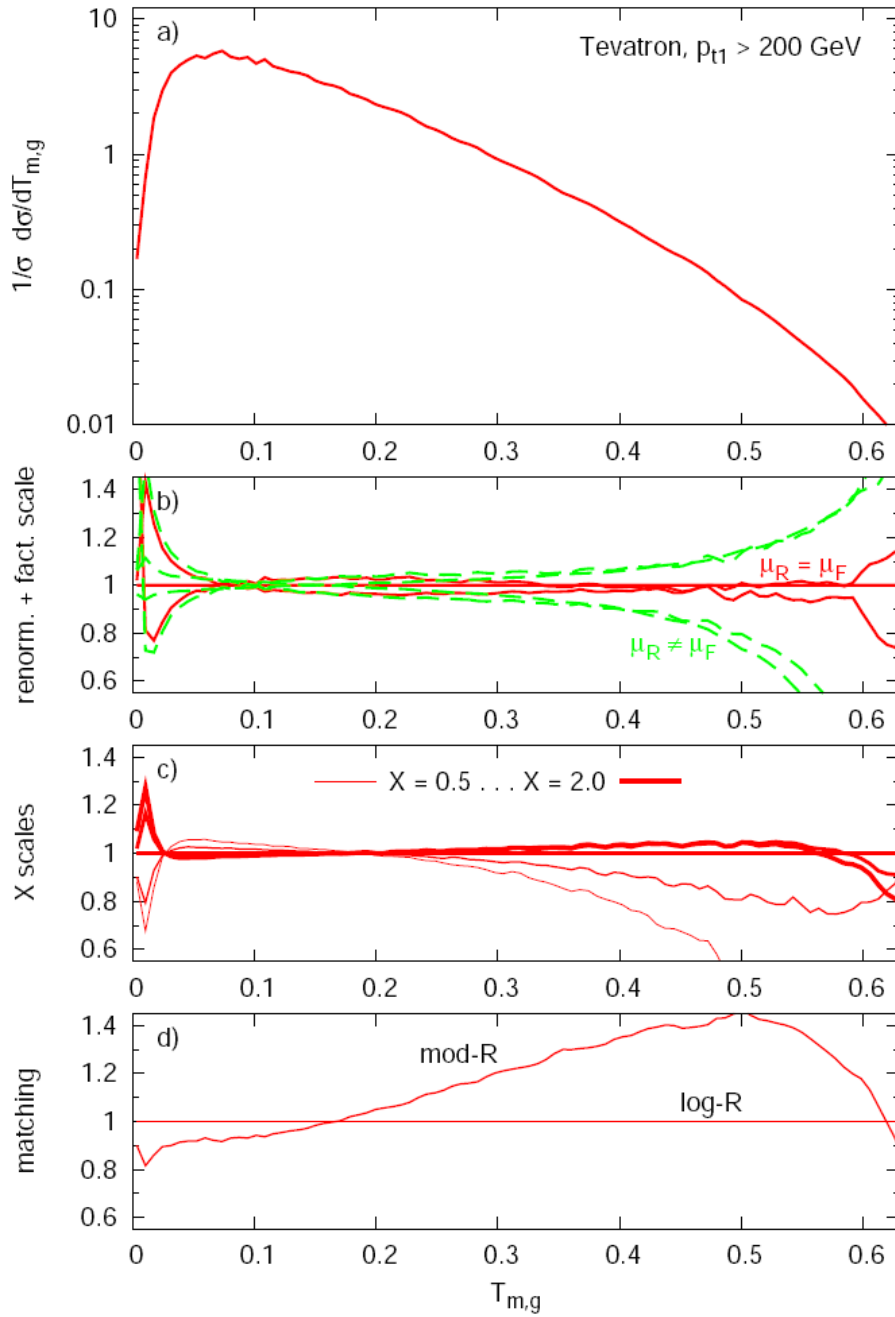
Figure 8.10: Resummed and matched results for $T_{m,g}$ at NLL+NLO accuracy as obtained in [3], from where the figure is extracted. The p_{t1}-cut is 200 GeV. a) distribution of $T_{m,g}$; b) uncertainties from μ_r and μ_f variations by factors of 2 and 1/2; c) effect of choosing $X = 0.5, 0.7, 1.0, 1.5, 2.0$; d) dependence on the choice of matching scheme.

8.5 Comparing results of Sherpa and Caesar

In this section, we compare the results obtained by SHERPA to the predictions obtained in [3] by CAESAR. For both we depict an estimate for the uncertainties by an error band. This band contains in case of CAESAR uncertainties from both neglected higher order contributions and systematics, because for CAESAR the latter ones are not negligible. In case of SHERPA, we found the systematics to be much smaller than the scale uncertainties. Hence, we will restrict our considerations to the latter. For both, the error bands are found by choosing for each bin the minimal and maximal value of the considered normalized histograms. The depicted curve itself corresponds to the default choices, which are for SHERPA $N_{\mathrm{max}} = 3$ and $Q_{\mathrm{cut}} = 30$ GeV as well as $\mu_{f,0}$ and $\mu_{r,0}$ as defined in section 8.3. For CAESAR, the default choices are $\mu_f^2 = \mu_r^2 = \vec{p}_t^2$ and log-R matching with $X = 1$.[2]

As a side remark to the SHERPA results: while the curve shows the $\mu_r^2 = \mu_{r,0}^2$ case, we can easily recover the $\mu_r^2 = 4\mu_{r,0}^2$ and the $\mu_r^2 = \mu_{r,0}^2/4$ curves from the uncertainty band. As explained in section 8.3, increasing μ_r will always shift the distribution to the left. Thus, before the narrow area close to the peak, where the curves of different μ_r values intersect, the $\mu_r^2 = 4\mu_{r,0}^2$ sample will correspond to the maximal value of the error band, while it will correspond to the minimal value behind that area. For the $\mu_r^2 = \mu_{r,0}^2/4$ sample it is the other way around.

In Figures 8.11 and 8.12, the plots for all observables as defined in sections 7.2 and the last part of 7.3 are given for $p_{t1,\mathrm{cut}} = 50$ GeV. These are the observables, which are either directly global or have exponentially suppressed forward terms. The central observables cannot be resummed in CAESAR, since they are not global. The observables with recoil terms have with very few exceptions huge error bands in the CAESAR sample and seem to be numerical unstable. Therefore, we will not show them here. In SHERPA, both variants give quite nice and stable results, which are similar to those of the other variants of the observables.

For all observables, we find that within their uncertainties the predictions from SHERPA and CAESAR agree. The estimated uncertainties are of comparable size. With the exception of the observables $y_{23,g/E}$ and $\rho_{H,E}$, the peak of the default curve from CAESAR is located at smaller values of the observable compared to the peak of the default curve from SHERPA. As a consequence, the latter is located below it before the peak and above it after the peak. Recall from the section about scale uncertainties in SHERPA that an effect like this can be obtained by varying μ_r and that it is the $\mu_r = 2\mu_{r,0}$ curve, which produces the maximum of the error band before the peak and the minimum after it. In fact, with the exception of y_{23}, out of the SHERPA samples always the $\mu_r = 2\mu_{r,0}$ sample agrees best with the CAESAR default curve. To this interesting point, we will come back in the end of this section.

Now, let us look at the corresponding results for the high $p_{t1,\mathrm{cut}}$ sample with $p_{t1,\mathrm{cut}} = 200$ GeV, which are shown in Figures 8.13 and 8.14. The observations correspond to the $p_{t1,\mathrm{cut}} = 50$ GeV case: Within the error range, the two predictions agree. Exceptions are

---

[2]As we have seen in chapter 6, in the resummed function $\vec{q}_\perp^2$ was chosen as renormalization scale. However, using the running of the coupling, the scale there can be changed to $\vec{p}_t^2$.
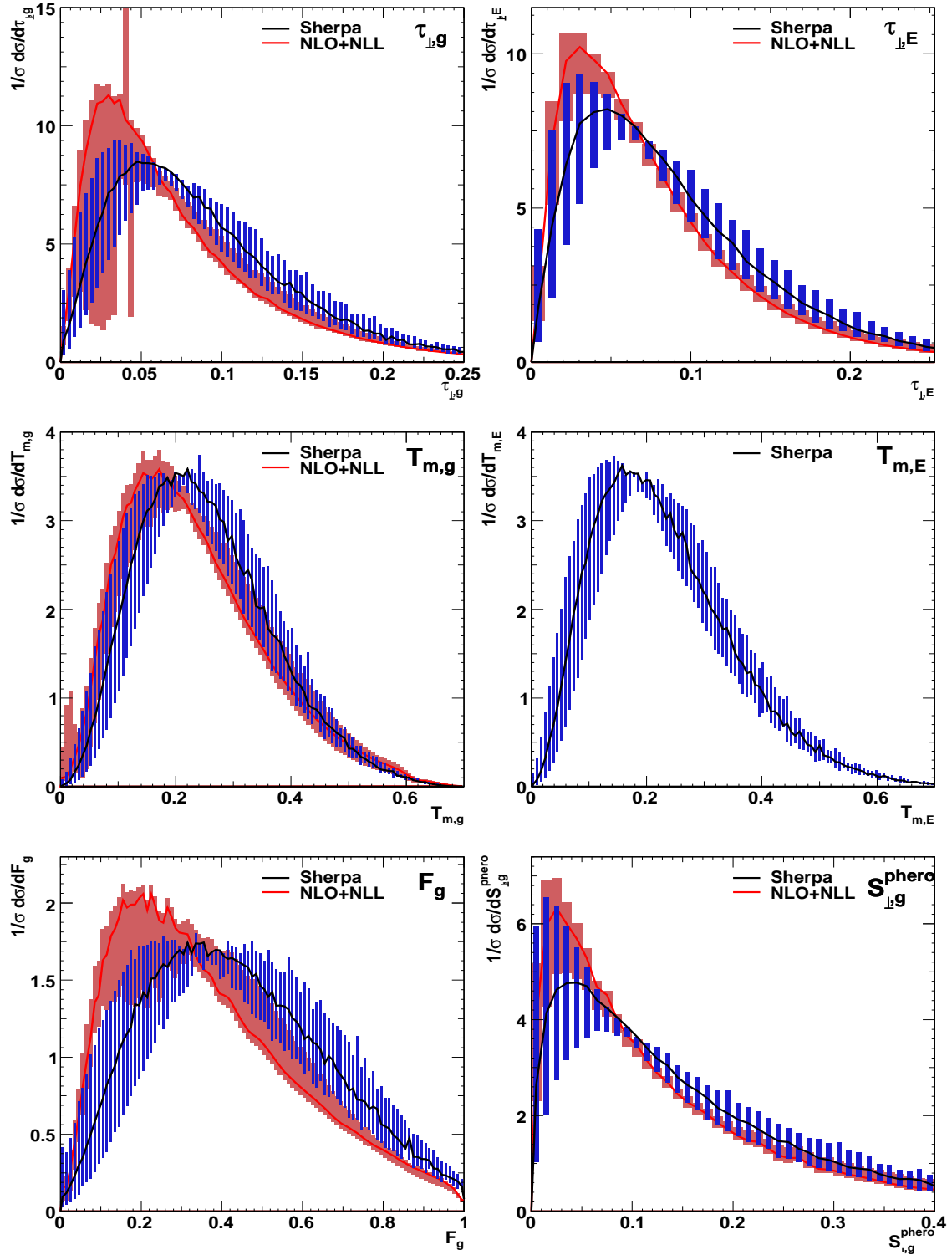
Figure 8.11: Exponentially suppressed and global event shape distributions for the 1.96 TeV Tevatron with $|\eta_c|=1$ and the cuts $p_{t1} > 50$ GeV and $|y_j| < 0.7$. Predictions of SHERPA (blue, foreground) are compared to predictions obtained with CAESAR (red, background). The bands represent the uncertainties.
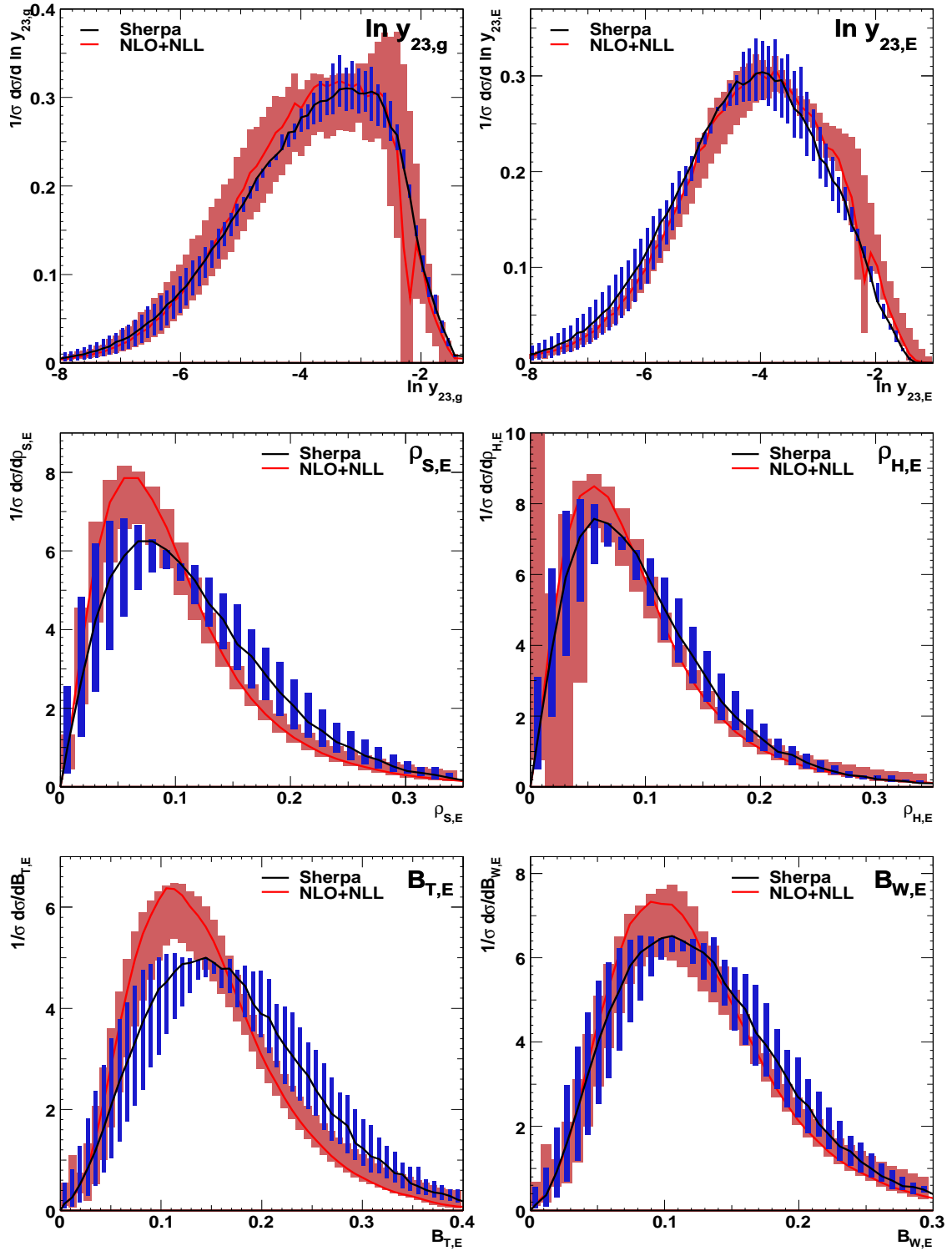
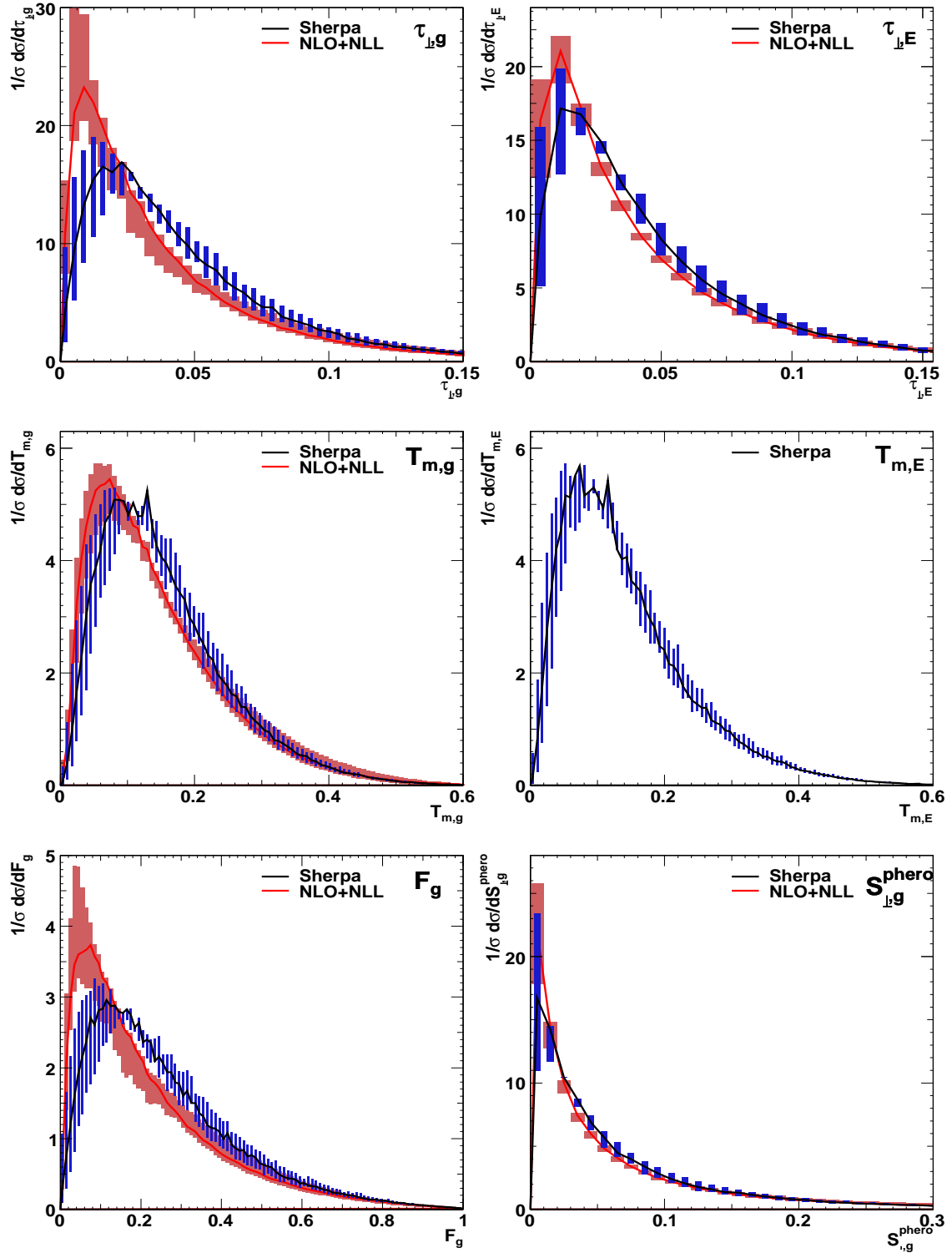Figure 8.12: Like Figure 8.11, but for jet observables.

Figure 8.13: Global and exponentially suppressed event shape distributions for the 1.96 TeV Tevatron with $|\eta_c|=1$ and the cuts $p_{t1} > 200$ GeV and $|y_j| < 0.7$. Predictions of SHERPA (blue, foreground) are compared to predictions obtained with CAESAR (red, background). The bands represent the uncertainties.
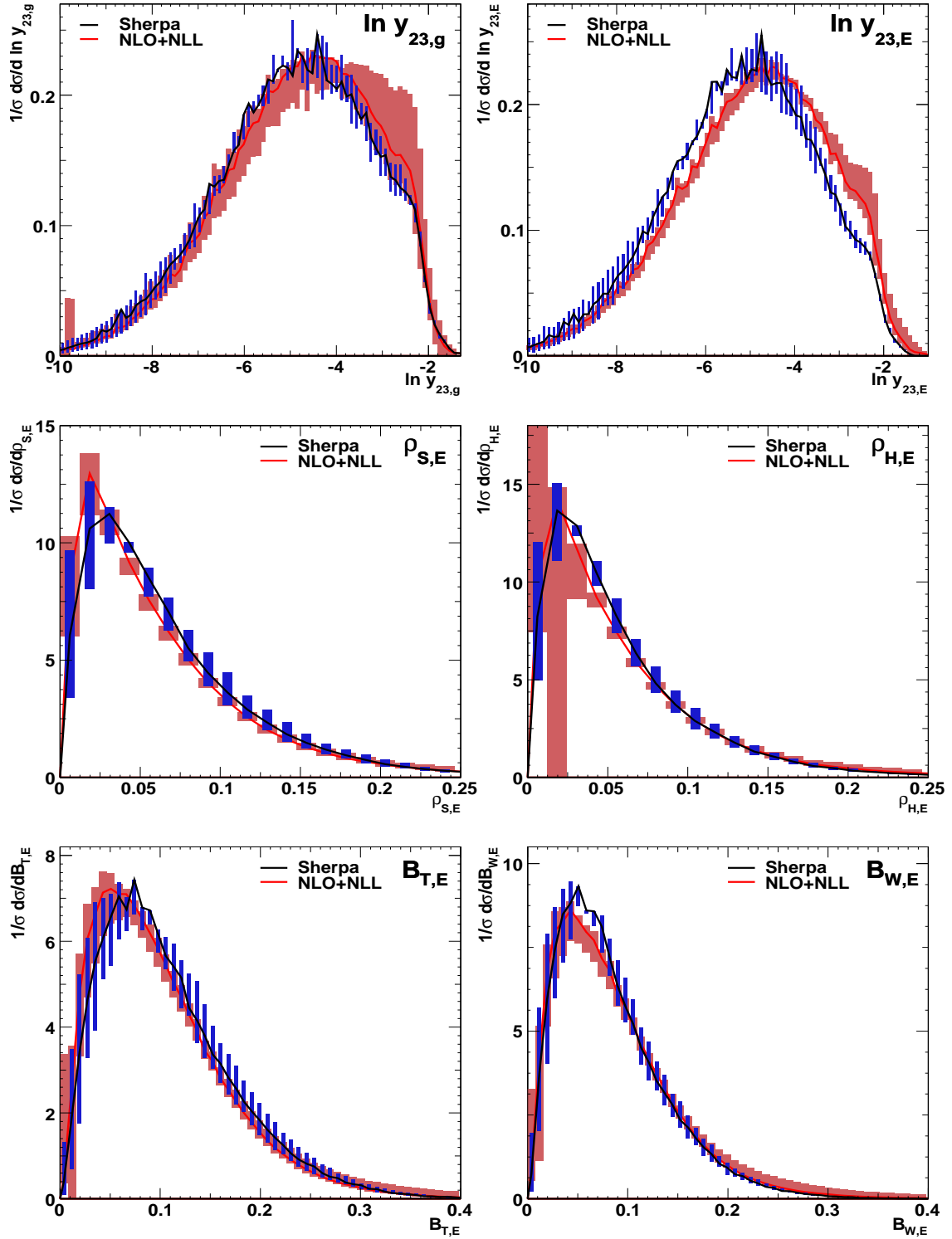
Figure 8.14: Like Figure 8.13, but for jet observables.

τ_g and F_g, whose peaks are that sharp for the CAESAR sample that in their region the error ranges do not overlap. The default CAESAR curve peaks left from the default SHERPA curve. The $\mu_r = 2\mu_{r,0}$ sample from SHERPA agrees best with the CAESAR samples with the exceptions of y_{23} and $B_{W,E}$.

For all observables, the comparison suggests that by choosing an appropriate value of μ_r, SHERPA can recover the event shape distribution as predicted by CAESAR within its uncertainties. For some observables, we may need to choose a different value for μ_r as for others. This is a very useful result. Although the resummation approach has the higher formal accuracy of NLO+NLL, it has the significant shortcoming that we do not know any way to include corrections from non-perturbative effects to it. Non-perturbative effects, for example, come from the underlying event and from hadronization. If we want to compare our results to experimental data, we, however, want to include corrections for such effects. At this point, the Monte Carlo parton shower programs are very useful. For them, we can include corrections like these straightforwardly by applying phenomenological models to describe the fragmentation of partons and the underlying event. Since those are applied after or in addition to the event generation on parton level, the parton shower and the merging are not affected by those. On parton level, we can adjust the program's settings by choice of μ_r to the predictions from resummation. For those settings, we can then additionally include corrections from hadronization and the underlying event. In this way, we probably obtain the most accurate predictions for the event shapes, which then can be compared to data.

In [3], the results from resummation, which we use here, were generated and compared to the Monte Carlo parton shower programs HERWIG 6.5 and PYTHIA 6.4. The predictions from HERWIG with the default choices of parameters are very similar to those from SHERPA. However, the estimated uncertainties from the symmetric variation of the renormalization and factorization scales are very small for HERWIG and obviously underestimate the real uncertainties, since the error band does not overlap with the error band of CAESAR. Similar small error bands were found for PYTHIA. Moreover, its default curve has worse agreement with the CAESAR predictions than SHERPA and HERWIG. Hence, of the three, SHERPA seems to perform best, especially because it leads to more realistic estimates of the uncertainties.

We showed in chapter 5 that in the large N_c limit the final-final Sudakov form factor in SHERPA has NLL accuracy in $\log Q^2/\vec{q}_{\perp,0}^2$ and supposed that this also holds for the three other Sudakov form factors. Thus, the parton shower in SHERPA will in the large N_c limit generate additional emissions with NLL accuracy. Because, in the relevant region of small values v of the observable, its value is directly related to the transverse momenta of the emissions, we expect that in this region the accuracy of the Sudakov form factor translates to the same accuracy of the event shape distribution in $\log 1/v$. Thus, the expected accuracy of the event shape's distributions from SHERPA is NLL in the large N_c limit. The agreement of the SHERPA and CAESAR samples within their uncertainties also supports this. To tighten this expectation, we should, however, provide an argument for the systematic shift of the SHERPA sample to larger values of μ_r, which leads in case of F_g and $\tau_{\perp,g}$ for $p_{t1,\text{cut}} = 200$ GeV to some deviation.

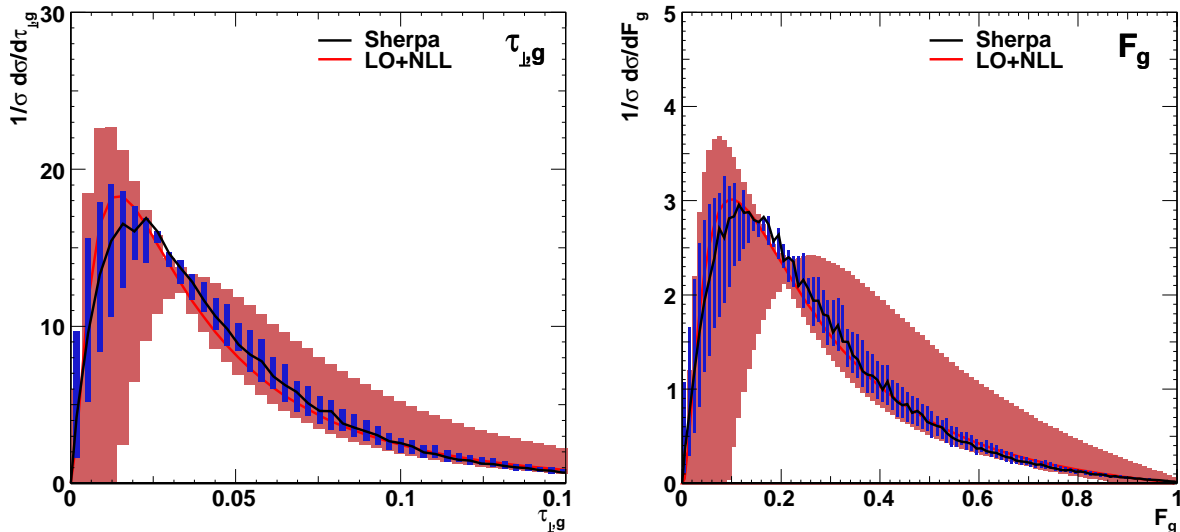The two most likely reasons are related to the higher formal precision of the CAESAR

Figure 8.15: $\tau_{\perp,g}$ and F_g for $p_{t1} > 200$ GeV and $|y_j| < 0.7$. Predictions of SHERPA (blue, foreground) are compared to the LO+NLL predictions from CAESAR (red, background). The bands represent the uncertainties.

prediction. On the one hand, CAESAR accounts for non-leading $1/N_c$ terms, which contribute NLL terms, while SHERPA neglects these terms. These might have the effect to shift the peak of the distribution to the left if, for example, the corresponding additional terms suppress radiation of hard partons systematically. On the other hand, CAESAR provides NLO precision in the differential cross section of v, while SHERPA only provides LO accuracy.[3] The NLO correction includes, the one-loop $2 \rightarrow 3$ terms and the tree-level $2 \rightarrow 4$ process. If these terms systematically reduce hard emission, this explains the shift of the distributions. The two emissions in the tree-level $2 \rightarrow 4$ process might interfere destructively and thus suppress the second emission. However, for this tree-level process SHERPA can account by the choice $N_{\max} \geq 4$ and we did not observe a strong dependence on the choice of $N_{\max} > 2$. Therefore, it should rather be the one-loop $2 \rightarrow 3$ terms, which suppress hard emissions and thus shift the peak of the CAESAR distribution to the left. In fact, the LO+NLL results from CAESAR are shifted to the right with respect to the NLO+NLL predictions and when they are compared to the SHERPA sample, no deviation is observed for F_g and $\tau_{\perp,g}$, as can be seen in Figure 8.5. We, therefore, think that the systematic shift between the two samples is caused by the inclusion of the NLO term.

To further investigate this point, one should vary the renormalization scales as they appear in the hard matrix element and in the shower of SHERPA separately. In this way, one gets a tool to estimate the expected higher order corrections for both the fixed order expansion and the resummed logarithms separately. Unfortunately, this is beyond the scope of this work.

---

[3] The LO term in the differential cross section of v is the tree-level $2 \rightarrow 3$ process, because the observable vanishes for the $2 \rightarrow 2$ process. Check footnote 2 on page 64.

For some observables the SHERPA prediction agrees even better with the NLO+NLL result than the LO+NLL result does. The reason for this should be that the emissions, as generated by the shower, give a good approximation to the full matrix element in this case.

The comparison of the SHERPA and CAESAR predictions, therefore, suggests that SHERPA has, in the large N_c limit, NLL accuracy. Moreover, for a comparison to experimental data the use of SHERPA with its included models for hadronization and underlying event is a preferable choice.

9 Conclusions

Starting from perturbative QCD, we reviewed general properties of QCD matrix elements and the multi particle phase space, which allowed for the factorization of soft and collinear radiation. We saw that this kind of radiation is logarithmically enhanced and introduced two approaches to take account for its dominant effects. On the one hand, the Monte Carlo parton shower of SHERPA, which uses a Monte Carlo method to extend fixed order calculations by explicitly generating additional splittings; on the other hand, the method of resummation of CAESAR, which follows a semi-analytical approach to account for logarithmic corrections by determination of an observable specific functions.

While the latter can provide the higher formal accuracy, the parton shower is much more flexible. It can be run independently of the specific choice of observables and detector cuts. Moreover, it offers possibilities to include further effects, such as hadronization, the underlying event and detector simulations. It, however, has a limited accuracy. To allow for the probabilistic and process independent picture of a parton shower, it has to ignore subleading color correlations and has to assume some kind of ordering among the emissions.

The approach of resummation of CAESAR provides the higher accuracy, as it does not neglect subleading color contributions and can be matched to fixed order matrix elements of higher precision, than a shower can be. This method, however, also has its shortcomings. It can only be applied to a limited range of observables fulfilling specific requirements. The calculation has to be performed separately for each observable and choice of detector cuts. For typical detector cuts as a pseudo-rapidity cut, it looses its formal accuracy. No way is know to include corrections from non-perturbative effects as e.g. hadronization to it.

Hence, a parton shower simulation and a resummed calculation offer complementary approaches to the same physics.

In our work, we studied the logarithmic accuracy of SHERPA. In case of a dipole of two massless final-state partons, we analytically determined the anomalous Sudakov dimensions and found them to be correct to NLL accuracy. We argued that this also holds for the other dipoles, which include initial-state partons. We also derived an expression for a rescaling of the renormalization scale, which allowed to include all other LL and NLL effects, which are leading in $1/N_C$, in the Sudakov form factor. Hence, this is the accuracy, which can be provided by SHERPA for the generation of events and the predictions of observables. From the discussion of CAESAR we learned, that there is a NLL term, which is non-leading in $1/N_c$. This term cannot be provided by SHERPA. There is also no hope to get potential NNLL terms right, as the evolution in SHERPA uses an ordering in the transverse momentum of the emitted particles, which is not correct at level of NNLL anymore. This shows the limitation of the parton shower approach.

In addition to considerations regarding formal accuracies, we performed a numerical

study to compare the predictions of SHERPA and CAESAR. To this end we considered the distributions of a range of dijet event shape variables for the Tevatron collider. These are observables, which measure the shape of the momentum flow in a scattering event. For pure QCD events, we mostly see events with only a few collimated jets, while for many new physics models, jets are expected to have a more spherical distribution. Hence, event shapes might be used in searches for new physics to improve the signal to background ratio.

Moreover, event shapes are useful to test perturbative QCD, because depending on the considered part of their distribution, the latter is described mostly by fixed order calculations or by the effect of successive parton splitting.

In this work, we were interested in their application to test and compare the perturbative QCD methods contained in SHERPA and CAESAR. We explained the general form of the event shape distributions. For SHERPA, we performed a detailed study of various uncertainties. The uncertainties introduced by the matching to multi-leg matrix elements were found to be small, which shows the consistency of the matching method used. Also the dependence on the factorization scale turned out to be weak. The dominant source of uncertainty was found to be the variation of the renormalization scale. Since SHERPA provides only LO accuracy in its fixed order part, sizeable corrections from higher orders and hence a strong dependence on the renormalization scale are expected.

The predictions from SHERPA were compared to predictions obtained with CAESAR in [3], which provide NLO+NLL accuracy. We found both predictions to agree within their uncertainties. However, we also observed a systematic shift of the peak of the SHERPA distribution to the right, which corresponds to a more frequent occurrence of hard emissions. Since this shift also appeared in the CAESAR sample with LO+NLL accuracy, we think that it is dominantly caused by NLO effects (in the differential cross section of the observable), which seem to suppress the occurrence of hard emissions. This shift could mostly be reproduced by SHERPA by increasing the renormalization scale by a factor of two.

This is a useful observation, because for a comparison of results to experimental data, one also needs to include corrections from effects, which cannot be described in CAESAR, such as hadronization or the underlying events. Since SHERPA can include corrections of this kind, it is the preferable choice for a direct comparison to experiments. With the adjusted renormalization scale, we know, that the parton level predictions correspond within uncertainties to the ones of NLO+NLL accuracy. In this way, we can combine the strengths of the two approaches for obtaining predictions, which will be phenomenologically relevant.

There are various points, which need further considerations. On the one hand, one should try to tell apart the uncertainties from the fixed order matrix element and the subsequent shower by varying the renormalization scales independently for both. On the other hand, one should also look at the distributions for other experiments. In especially to those of the LHC, where they might be relevant or searches for new physics.

Acknowledgements

9 Conclusions

Bibliography

[1] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009), 007, [`arXiv:0811.4622` [hep-ph]].

[2] A. Banfi, G. P. Salam and G. Zanderighi, *Principles of general final-state resummation and automated implementation*, JHEP **03** (2005), 073, [`hep-ph/0407286`].

[3] A. Banfi, G. Salam and G. Zanderighi, *Phenomenology of event shapes at hadron colliders*, JHEP **06** (2010), 038, [`arXiv:1001.4082` [hep-ph]].

[4] C. A. et al. (Particle Data Group), *The Review of Particle Physics*, Physics Letters **B667** (2008), no. 1

[5] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*, Westview, 1995.

[6] S. Weinberg, *The Quantum Theory of Fields*, I-III, 1996.

[7] L. Faddeev and V. Popov, Phys. Lett. **25B** (1967), no. 29

[8] K. Ellis, W. Stirling and B. Webber, *QCD and Collider Physics*, Cambridge University Press, 1996.

[9] T. Kinoshita, *Mass singularities of Feynman amplitudes*, J. Math. Phys. **3** (1962), 650–677.

[10] T. Lee and M. Nauenberg, *Degenerate Systems and Mass Singularities*, Phys. Rev. **133** (1964), B1549–B1562.

[11] G. F. Sterman, *Some Basic Concepts of Perturbative QCD*, Acta Phys. Polon. **B39** (2008), 2151–2172, [`arXiv:0807.5118` [hep-ph]].

[12] T. Ritbergen, J. Vermaseren and S. Larin, *The four-loop beta function in quantum chromodynamics*, Phys. Lett. **B400** (1997), 379–384.

[13] J. Bjorken, Phys. Rev. **179** (1969), 1547.

[14] S.-M. collaboration collaboration collaboration collaboration, Phys. Rev. **D5** (1972), 528.

[15] J. Collins, D. Soper and G. Sterman, *Factorization of Hard Processes in QCD*, Adv. Ser. Direct. High Energy Phys. **5** (1988), 1–91, [`hep-ph/0409313`].

Bibliography

[16] H. Politizer, Phys. Rept. **14** (1974), no. 129

[17] D. Gross and F. Wilczek, Phys. Rev. **D8** (1973), no. 3633

[18] G. Bodwin, Phys. Rev. **D31** (1985), no. 2616, 104, Erratum ibid. D3, 3932 (1986).

[19] G. Altarelli and G. Parisi, Nucl. Phys. **B126** (1977), 298.

[20] V. Gribov and L. Lipatov, Sov.J.Nucl.Phys **15** (1972), 438.

[21] T. Plehn, *An LHC Lecture*, `arXiv:0910.4182` [hep-ph].

[22] S. Catani, F. Krauss, R. Kuhn and B. Webber, *QCD Matrix Elements + Parton Showers*, JHEP **11** (2001), 063, [`hep-ph/0109231`].

[23] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008), 038, [`arXiv:0709.1027` [hep-ph]].

[24] S. Catani and M. Seymour, *A general algorithm for calculating jet cross sections in NLO QCD*, Nucl. Phys. **B485** (1997), 291–419, [`hep-ph/9605323`].

[25] S. Hoeche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **05** (2009), 053, [`arXiv:0903.1219` [hep-ph]].

[26] F. Krauss, R. Kuhn and G. Soff, *AMEGIC++ 1.0: A Matrix element generator in C++*, JHEP **02** (2002), 044, [`hep-ph/0109036`].

[27] T. Gleisberg and S. Hoeche, *Comix, a new matrix element generator*, JHEP **12** (2008), 039, [`arXiv:0808.3674` [hep-ph]].

[28] T. Sjostrand and M. van Zijl, *A Multiple Interaction Model for the Event Structure in Hadron Collisions*, Phys. Rev. **D36** (1987), 2019.

[29] J.-C. Winter, F. Krauss and G. Soff, *A modified cluster-hadronization model*, Eur. Phys. J. **C36** (2004), 381–395, [`hep-ph/0311085`].

[30] M. Schonherr and F. Krauss, *Soft Photon Radiation in Particle Decays in SHERPA*, JHEP **12** (2008), 018, [`arXiv:0810.5071` [hep-ph]].

[31] S. Catani, S. Dittmaier, M. Seymour and Z. Trocsanyi, *The dipole formalism for next-to-leading order QCD calculations with massive partons*, Nucl. Phys. **B627** (2002), 189–265, [`hep-ph/0201036`].

[32] S. Plaetzer and S. Gieseke, *Coherent Parton Showers with Local Recoils*, `arXiv:0909.5593` [hep-ph].

[33] S. Catani, B. Webber and G. Marchesini, *QCD coherent branching and semiinclusive processes at large x*, Nucl. Phys. **B349** (1991), 635–654.

[34] S. Catani, L. Trentadue, G. Turnock and B. R. Webber, *Resummation of large loga-rithms in e+ e- event shape distributions*, Nucl. Phys. **B407** (1993), 3–42.

[35] M. Dasgupta and G. Salam, *Resummed event-shape variables in DIS*, JHEP **08** (2002), 032, [`hep-ph/0208073`].

[36] Z. Nagy, *Next-to-leading order calculation of three jet observables in hadron hadron collision*, Phys. Rev. **D68** (2003), 094002, [`hep-ph/0307268`].

[37] A. Banfi, G. Salam and G. Zanderighi, *Resummed event shapes at hadron - hadron colliders*, JHEP **08** (2004), 062, [`hep-ph/0407287`].

[38] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, *Longitudinally invari-ant K_t clustering algorithms for hadron hadron collisions*, Nucl. Phys. **B406** (1993), 187–224.

[39] A. Banfi, G. Marchesini, G. Smye and G. Zanderighi, *Out-of-plane QCD radiation in hadronic Z0 production*, JHEP **08** (2001), 047, [`hep-ph/0106278`].

[40] G. P. Salam and G. Soyez, *A practical Seedless Infrared-Safe Cone jet algorithm*, JHEP **05** (2007), 086, [`arXiv:0704.0292` [hep-ph]].

Bibliography

Erklärung:

Ich versichere, dass ich diese Arbeit selbstständig verfasst habe und keine anderen als die
angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 11.10.2010   . . . .