Elements of QCD for hadron colliders

G. P. Salam
CERN, Department of Theoretical Physics, Geneva, Switzerland
University of Oxford, Oxford, United Kingdom

Abstract
The aim of these lectures is to provide students with an introduction to some of the core concepts and methods of QCD that are relevant in an LHC context.

1 Introduction

Quantum Chromodynamics, QCD, is the theory of quarks, gluons and their interactions. It is central to all modern colliders. And, for the most part, it is what we are made of.

QCD bears a number of similarities to Quantum Electrodynamics (QED). Just as electrons carry the QED charge, i.e., electric charge, quarks carry the QCD charge, known as colour charge. Whereas there is only one kind of electric charge, colour charge comes in three varieties, sometimes labelled red, green and blue. Anti-quarks have corresponding anti-colour. The gluons in QCD are a bit like the photons of QED. But while photons are electrically neutral, gluons are not colour neutral. They can be thought of as carrying both colour charge and anti-colour charge. There are eight possible different combinations of (anti)colour for gluons. Another difference between QCD and QED lies in its coupling α_s. In QCD it is only moderately small, it tends to zero at high momentum scales (asymptotic freedom, QED does the opposite), it blows up at small scales, and in between its evolution with scale is quite fast: at the LHC its value will range from $\alpha_s = 0.08$ at a scale of $5\,\text{TeV}$, to $\alpha_s \sim 1$ at a scale of $0.5\,\text{GeV}$. These differences between QCD and QED contribute to making QCD a much richer theory.

In these lectures I will attempt to give you a feel for how QCD works at high momentum scales, and for the variety of techniques used by theorists in order to handle QCD at today's high-energy colliders. The hope is that these basics will come in useful for day-to-day work with the QCD facets of hadron collider physics. In the fifty or so pages of these lectures, it will be impossible to give full treatment of any of the topics we will encounter. For that the reader is referred to any of the classic textbooks about QCD at colliders [1–3].

1.1 The Lagrangian and colour

Let us start with a brief reminder of the components of the QCD Lagrangian. This section will be rather dense, but we will return to some of the points in more detail later. As already mentioned, quarks come in three colours. So rather than representing them with a single spinor ψ, we will need the spinor to carry also a colour index a, which runs from 1 to 3,

$$\psi_a = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix}.\tag{1}$$

The quark part of the Lagrangian (for a single flavour) can be written

$$\mathcal{L}_q = \bar{\psi}_a(i\gamma^\mu \partial_\mu \delta_{ab} - g_s \gamma^\mu t^C_{ab} \mathcal{A}^C_\mu - m)\psi_b\,,\tag{2}$$

where the γ^μ are the usual Dirac matrices; the $\mathcal{A}_\mu^C$ are gluon fields, with a Lorentz index μ and a colour index C that goes from $1 \ldots 8$. Quarks are in the fundamental representation of the SU(3) (colour) group, while gluons are in the adjoint representation. Each of the eight gluon fields acts on the quark colour through one of the 'generator' matrices of the SU(3) group, the t_{ab}^C factor in Eq. (2). One convention for writing the matrices is $t^A \equiv \frac{1}{2}\lambda^A$ with

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; \lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$\lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \; \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \; \lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \; \lambda^8 = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 \\ 0 & 0 & \frac{-2}{\sqrt{3}} \end{pmatrix}.$$

By looking at the first of these, together with the $t_{ab}^C \mathcal{A}_\mu^C \psi_b$ term of $\mathcal{L}_Q$, one can immediately get a feel for what gluons do: a gluon with (adjoint) colour index $C = 1$ acts on quarks through the matrix $t^1 = \frac{1}{2}\lambda^1$. That matrix takes green quarks ($b = 2$) and turns them into red quarks ($a = 1$), and vice versa. In other words, when a gluon interacts with a quark it *repaints* the colour of the quark, taking away one colour and replacing it with another. The likelihood with which this happens is governed by the strong coupling constant g_s. Note that the repainting analogy is less evident for some of the other colour matrices, but it still remains essentially correct.

The second part of the QCD Lagrangian is purely gluonic

$$\mathcal{L}_G = -\frac{1}{4} F_A^{\mu\nu} F^{A\,\mu\nu} \tag{3}$$

where the gluon field tensor $F_{\mu\nu}^A$ is given by

$$F_{\mu\nu}^A = \partial_\mu \mathcal{A}_\nu^A - \partial_\nu \mathcal{A}_\mu^A - g_s\, f_{ABC}\mathcal{A}_\mu^B \mathcal{A}_\nu^C \qquad [t^A, t^B] = i f_{ABC} t^C\,, \tag{4}$$

where the f_{ABC} are the structure constants of SU(3) (defined through the commutators of the t^A matrices). Note the major difference with QED here, namely the presence of a term $g_s\, f_{ABC}\mathcal{A}_\mu^B \mathcal{A}_\nu^C$ with two gluon fields. The presence of such a term is one of the major differences with QED, and, as we will discuss in more detail below, it will be responsible for the fact that gluons interact directly with gluons. For now, note simply that it has to be there in order for the theory to be gauge invariant under local SU(3) transformations:

$$\psi_a \to\; e^{i\theta_C(x) t_{ab}^C}\psi_b \tag{5}$$

$$\mathcal{A}^C t^C \to\; e^{i\theta^D(x) t^D}\left(\mathcal{A}^C t^C - \frac{1}{g_s}\partial_\mu \theta^C(x) t^C\right) e^{-i\theta^E(x) t^E} \tag{6}$$

where, in the second line, we have dropped the explicit subscript ab indices, and the $\theta^C(x)$ are eight arbitrary real functions of the space-time position x.

1.2 'Solving QCD'

There are two main first-principles approaches to solving QCD: lattice QCD and perturbative QCD.[1]

---

[1] In addition, effective-theory methods provide ways of looking at QCD that make it easier to solve, given certain 'inputs' that generally come from lattice or perturbative QCD (and sometimes also from experimental measurements). These lectures won't discuss effective theory methods, but for more details you may consult the lectures at this school by Martin Beneke. Another set of methods that has seen much development in recent years makes use of the 'AdS/CFT' correspondence [4–6], relating QCD-like models at strong coupling to gravitational models at weak coupling (e.g., [7,8]).
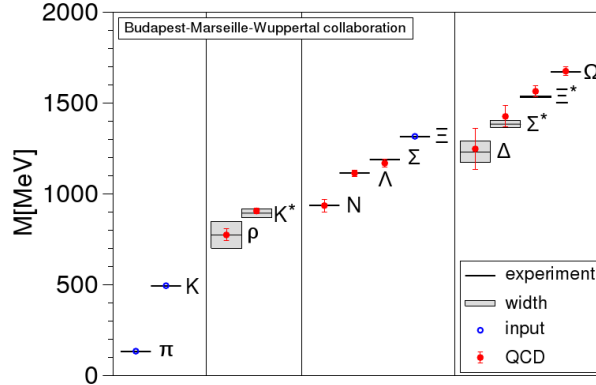
Fig. 1: The measured spectrum of hadron masses, compared to a lattice calculation [9]. The open blue circles are the hadron masses that have been used to fix the three parameters of the calculation: the value of the QCD coupling, the average of the up and down quark masses (taken equal) and the strange-quark mass. All other points are results of the calculation.

1.2.1 Lattice QCD

The most complete approach is lattice QCD. It involves discretizing space-time, and considering the values of the quark and gluon fields at all the vertices/edges of the resulting 4-dimensional lattice (with imaginary time). Through a suitable Monte Carlo sampling over all possible field configurations, one essentially determines the relative likelihood of different field configurations, and this provides a solution to QCD. This method is particularly suited to the calculation of static quantities in QCD such as the hadron mass spectrum. The results of such a lattice calculation are illustrated in Fig. 1, showing very good agreement.

Lattice methods have been successfully used in a range of contexts, for example, in recent years, in helping extract fundamental quantities such as the CKM matrix (and limits on new physics) from the vast array of experimental results on hadron decays and oscillations at flavour factories. Unfortunately lattice calculations aren't suitable in all contexts. Let us imagine, briefly, what would be required in order to carry out lattice calculations for LHC physics: since the centre-of-mass energy is (will be) $14 \, \text{TeV}$, we need a lattice spacing of order $1/(14 \, \text{TeV}) \sim 10^{-5} \, \text{fm}$ to resolve everything that happens. Non-perturbative dynamics for quarks/hadrons near rest takes place on a timescale $t \sim \frac{1}{0.5 \, \text{GeV}} \sim 0.4 \, \text{fm}/c$. But hadrons at LHC have a boost factor of up to 10^4, so the extent of the lattice should be about $4000 \, \text{fm}$. That tells us that if we are to resolve high-momentum transfer interactions and at the same time follow the evolution of quark and gluon fields up to the point where they form hadrons, we would need about 4×10^8 lattice units in each direction, of $\sim 3 \times 10^{34}$ nodes. Not to mention the problem with high particle multiplicities (current lattice calculations seldom involve more than two or three particles) and all the issues that relate to the use of imaginary time in lattice calculations. Of course, that's not to say that it might not be possible, one day, to find clever tricks that would enable lattice calculations to deal with high-energy reactions. However, with today's methods, any lattice calculation of the properties of LHC proton–proton scattering seems highly unlikely. For this reason, we will not give any further discussion of lattice QCD here, but instead refer the curious reader to textbooks and reviews for more details [10–13].

1.2.2 Perturbative QCD

Perturbative QCD relies on the idea of an order-by-order expansion in a small coupling $\alpha_{\text{s}} = \frac{g_s^2}{4\pi} \ll 1$. Some given observable f can then be predicted as

$$f = f_1 \alpha_{\text{s}} + f_2 \alpha_{\text{s}}^2 + f_3 \alpha_{\text{s}}^3 + \dots , \qquad (7)$$
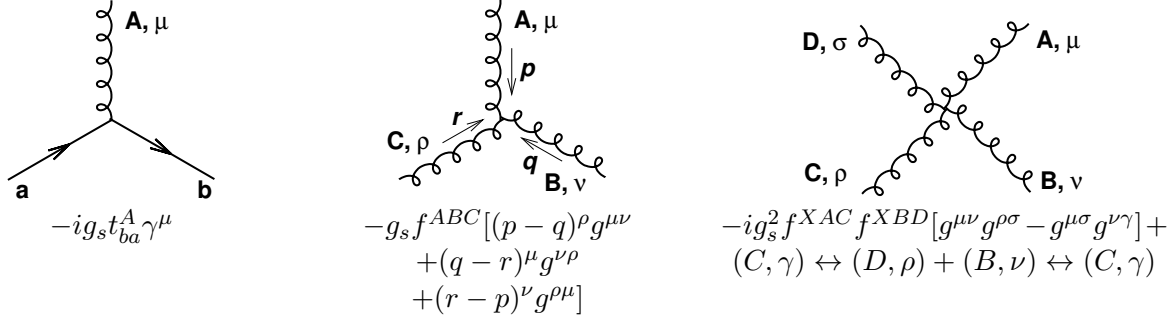
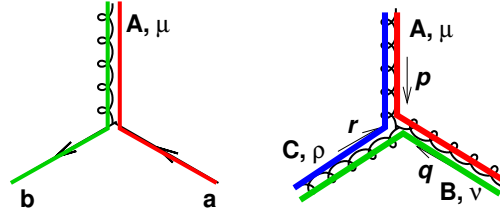**Fig. 2:** The interaction vertices of the Feynman rules of QCD

Fig. 3: Schematic colour flow interpretation of the quark–quark–gluon (t_{ab}^A, left) and triple-gluon (f_{ABC}, right) vertices of QCD. These interpretations are only sensible insofar as one imagines that the number of colours in QCD, $N_c = 3$, is large.

where one might calculate just the first one or two terms of the series, with the understanding that remaining ones should be small.
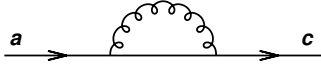
The principal technique to calculate the coefficients f_i of the above series is through the use of Feynman diagrammatic (or other related) techniques. The interaction vertices of the QCD Feynman rules are shown in Fig. 2 (in some gauges one also needs to consider ghosts, but they will be irrelevant for our discussions here).
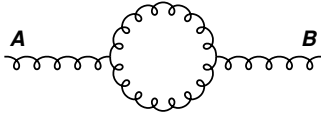
The qqg interaction in Fig. 2 comes from the $\bar{\psi}_a g_s \gamma^\mu t_{ab}^C \mathcal{A}_\mu^C \psi_b$ term of the Lagrangian. We have already discussed how the t_{ab}^C factor affects the colour of the quark, and this is represented in Fig. 3(left), with the gluon taking away one colour and replacing it with another.
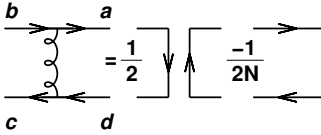
The triple-gluon vertex in Fig. 2 comes from the $-\frac{1}{4} F_A^{\mu\nu} F^{A\,\mu\nu}$ part of the Lagrangian, via the product of a $\partial_\mu \mathcal{A}_\nu$ term in one $F_A^{\mu\nu}$ factor with the $g_s f_{ABC} \mathcal{A}_\mu^B \mathcal{A}_\nu^C$ term in the other. It is the fact that gluons carry colour charge that means that they must interact with other gluons. In terms of colour flows, we have the repetition of the idea that the emission of a gluon can be seen as taking away the colour from the gluon (or anti-colour) and replacing it with a different one. Because of the double colour/anti-colour charge of a gluon, one can anticipate that it will interact with (or emit) other gluons twice as strongly as does a quark. Before coming to mathematical formulation of that statement, let's comment also on the 4-gluon vertex of Fig. 2. This comes from the product of two $g_s f_{ABC} \mathcal{A}_\mu^B \mathcal{A}_\nu^C$ type terms in $-\frac{1}{4} F_A^{\mu\nu} F^{A\,\mu\nu}$ and is order g_s^2 whereas the two other interactions are order g_s.

Though Fig. 3 gives some idea of how the colour factors t_{ab}^C and f_{ABC} in the Feynman rules are to be understood, it is useful to see also how they arise in calculations. After squaring an amplitude and summing over colours of incoming and outgoing particles, they often appear in one or other of the following combinations:

$$\mathrm{Tr}(t^A t^B) = T_R \delta^{AB}, \quad T_R = \frac{1}{2} \qquad \text{(8a)}$$

$$\sum_A t_{ab}^A t_{bc}^A = C_F \delta_{ac}, \quad C_F = \frac{N_C^2 - 1}{2N_C} = \frac{4}{3}$$ (8b)

$$\sum_{C,D} f^{ACD} f^{BCD} = C_A \delta^{AB}, \quad C_A = N_C = 3$$ (8c)

$$t_{ab}^A t_{cd}^A = \frac{1}{2} \delta_{bc} \delta_{ad} - \frac{1}{2N_C} \delta_{ab} \delta_{cd} \text{ (Fierz)}$$ (8d)

where $N \equiv N_C = 3$ is the number of colours in QCD and it is useful to express the results for general numbers of colours (because it is sometimes useful to consider how results depend on N_C, especially in the limit $N_C \to \infty$). Each mathematical combination of colour factors has a diagrammatic interpretation. Equation (8a) corresponds to a gluon splitting into $q\bar{q}$ which then join back into a gluon; or, the sum over colours in the squared amplitude for $g \to q\bar{q}$. Equation (8b) corresponds to the square of gluon emission from a quark. Equation (8c) arises as the square of gluon emission from a gluon. One sees that there is almost a factor of 2 between Eqs. (8b) and (8c) (modulo corrections terms $\sim 1/N_C$), which is the mathematical counterpart of our statement above that gluons emit twice as strongly as quarks. Finally the approximate colour-flow interpretation that we had in Fig. 3 (left) can be stated exactly in terms of the Fierz identity, Eq. (8d).

1.2.3   The running coupling

Most higher-order QCD calculations are carried out with dimensional regularization (the use of $4 - \epsilon$ dimensions) in order to handle the ultraviolet divergences that appear in loop diagrams. In the process of going from 4 to $4 - \epsilon$ dimensions, one needs to introduce an arbitrary 'renormalization' scale, generally called μ, in order to keep consistent dimensions (units) for all quantities.[2] The value of the QCD coupling, $\alpha_s = \frac{g_s^2}{4\pi}$, depends on the scale μ at which it is evaluated. That dependence can be expressed in terms of a renormalization group equation

$$\frac{d\alpha_s(\mu^2)}{d\ln\mu^2} = \beta(\alpha_s(\mu^2)), \qquad \beta(\alpha_s) = -\alpha_s^2(b_0 + b_1\alpha_s + b_2\alpha_s^2 + \dots),$$ (9)

where

$$b_0 = \frac{11C_A - 2n_f}{12\pi}, \qquad b_1 = \frac{17C_A^2 - 5C_A n_f - 3C_F n_f}{24\pi^2} = \frac{153 - 19n_f}{24\pi^2},$$ (10)

with n_f being the number of 'light' quark flavours, those whose mass is lower than μ. The negative sign in Eq. (9) is the origin of asymptotic freedom, the fact that the coupling becomes weaker at high momentum scales, i.e., the theory almost becomes a free theory, in which quarks and gluons don't interact. Conversely at low momentum scales the coupling grows strong, causing quarks and gluons to be tightly bound into hadrons. The importance of the discovery of these features was recognized in the 2004 Nobel prize to Gross, Politzer and Wilczek. Why does the QCD β-function have the opposite sign of that in QED? The fact that the vector particles (gluons) of the theory carry colour charge is central to the result. However, while there have been various attempts to give simple but accurate explanations for the negative sign [15, 16], in practice they all end up being quite involved.[3] So, for the purpose of these lectures, let us just accept the results.

---

[2]The renormalization procedure itself, i.e., the removal of the $1/\epsilon$ divergences, is usually carried out in the modified minimal subtraction ($\overline{\text{MS}}$) scheme (see, e.g., Section 11.4 of Ref. [14]), by far the most widespread scheme in QCD.

[3]You might still want to check the sign for yourself: if so, pick up a copy of Peskin and Schroeder [14], arrange to have an afternoon free of interruptions, and work through the derivation.
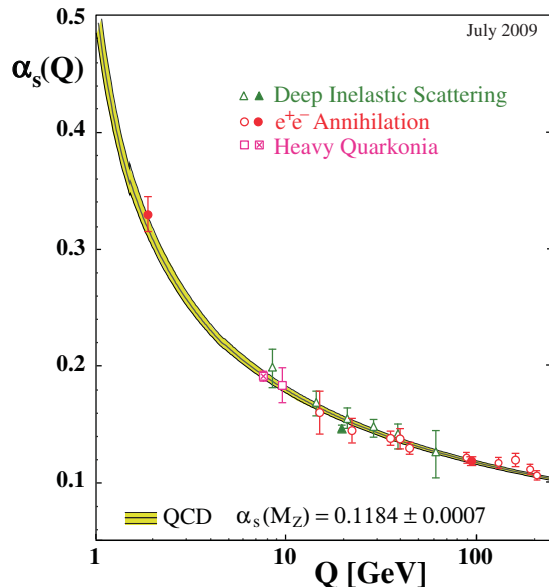
Fig. 4: The QCD coupling as measured in physics processes at different scales Q, together with the band obtained by running the world average for $\alpha_{\rm s}$ within its uncertainties. Figure taken from Ref. [17].

If we ignore all terms on the right of Eq. (9) other than b_0, and also ignore the subtlety that the number of 'light' flavours n_f depends on μ, then there is a simple solution for $\alpha_{\rm s}(\mu^2)$:

$$\alpha_{\rm s}(\mu^2) = \frac{\alpha_{\rm s}(\mu_0^2)}{1 + b_0 \alpha_{\rm s}(\mu_0^2) \ln \frac{\mu^2}{\mu_0^2}} = \frac{1}{b_0 \ln \frac{\mu^2}{\Lambda^2}} \,, \tag{11}$$

where one can either express the result in terms of the value of the coupling at a reference scale μ_0, or in terms of a non-perturbative constant Λ (also called Λ_{QCD}), the scale at which the coupling diverges. Only for scales $\mu \gg \Lambda$, corresponding to $\alpha_{\rm s}(\mu^2) \ll 1$, is perturbation theory valid. Note that Λ, since it is essentially a non-perturbative quantity, is not too well defined: for a given $\alpha_{\rm s}(\mu_0)$, its value depends on whether we used just b_0 in Eq. (9) or also b_1, etc. However, its order of magnitude, $200\,{\rm MeV}$, is physically meaningful insofar as it is closely connected with the scale of hadron masses.

One question that often arises is how μ, the renormalization scale, should relate to the physical scale of the process. We will discuss this in detail later (Section 4.1), but for now the following simple statement is good enough: the strength of the QCD interaction for a process involving a momentum transfer Q is given by $\alpha_{\rm s}(\mu)$ with $\mu \sim Q$. One can measure the strength of that interaction in a range of processes, at various scales, and Fig. 4 [17] shows a compilation of such measurements, together with the running of an average over many measurements, $\alpha_{\rm s}(M_Z) = 0.1184 \pm 0.0007$, illustrating the good consistency of the measurements with the expected running.

1.2.4 QCD predictions and colliders

Colliders like the Tevatron and the LHC are mainly geared to investigating phenomena involving high-momentum transfers (more precisely large transverse-momenta), say in the range $50\,{\rm GeV}$ to $5\,{\rm TeV}$. There, the QCD coupling is certainly small and we would hope to be able to apply perturbation theory. Yet, the initial state involves protons, at whose mass scale, $m_p \simeq 0.94\,{\rm GeV}$, the coupling is certainly not weak. And the final states of collider events consist of lots of hadrons. Those aren't perturbative either. And there are lots of them — tens to hundreds. Even if we wanted to try, somehow, to treat them perturbatively, we would be faced with calculations to some very high order in $\alpha_{\rm s}$, at least as high as the particle multiplicity, which is far beyond what we can calculate exactly: depending on how you count,

6

at hadron colliders, the best available complete calculation (i.e., all diagrams at a given order), doesn't go beyond α_s^2 or α_s^3. Certain subsets of diagrams (e.g., those without loops) can be calculated up α_s^{10} roughly.

So we are faced with a problem. Exact lattice methods can't deal with the high momentum scales that matter, exact perturbative methods can't deal with low momentum scales that inevitably enter the problem, nor the high multiplicities that events have in practice. Yet, it turns out that we are reasonably successful in making predictions for collider events. These lectures will try to give you an understanding of the methods and approximations that are used.

2 Considering $e^+e^- \to$ hadrons

One simple context in which QCD has been extensively studied over the past 30 years is that of e^+e^- annihilation to hadrons. This process has the theoretical advantage that only the final state involves QCD. Additionally, huge quantities of data have been collected at quite a number of colliders, including millions of events at the Z mass at LEP and SLC. We therefore start our investigation of the properties of QCD by considering this process.

2.1 Soft and collinear limits

There is one QCD approximation that we will repeatedly make use of, and that is the soft and collinear approximation. 'Soft' implies that an emitted gluon has very little energy compared to the parton (quark or gluon) that emitted it. 'Collinear' means that it is emitted very close in angle to another parton in the event. By considering gluons that are soft and/or collinear one can drastically simplify certain QCD calculations, while still retaining much of the physics.

The soft and collinear approximation is sufficiently important that it's worth, at least once, carrying out a calculation with it, and we'll do that in the context of the emission of a gluon from $e^+e^- \to q\bar{q}$ events. Though there are quite a few equations in the page that follows, the manipulations are all quite simple! We're interested in the hadronic side of the $e^+e^- \to q\bar{q}$ amplitude, so let's first write the QED matrix element for a virtual photon $\gamma^* \to q\bar{q}$ (we can always put back the $e^+e^- \to \gamma^*$ and the photon propagator parts later if we need to — which we won't):

$$\mathcal{M}_{q\bar{q}} = \bar{u}_a(p_1) i e_q \gamma_\mu \delta_{ab} v_b(p_2)$$

where the diagram illustrates the momentum labelling. Here $\bar{u}(p_1)$ and $v(p_2)$ are the spinors for the outgoing quark and anti-quark (taken massless), e_q is the quark's electric charge and the γ_μ are the Dirac matrices. In what follows we shall drop the a, b quark colour indices for compactness and reintroduce them only at the end.

The corresponding amplitude including the emission of a gluon with momentum k and polarization vector ϵ is

$$\mathcal{M}_{q\bar{q}g} = \quad\text{(diagram)}\quad + \quad\text{(diagram)}$$

$$= -\bar{u}(p_1) i g_s \not{\epsilon} t^A \frac{i(\not{p}_1 + \not{k})}{(p_1 + k)^2} i e_q \gamma_\mu v(p_2) + \bar{u}(p_1) i e_q \gamma_\mu \frac{i(\not{p}_2 + \not{k})}{(p_2 + k)^2} i g_s \not{\epsilon} t^A v(p_2) \,,$$

(12)

with one term for emission from the quark and the other for emission from the anti-quark and use of the notation $\not{p} = p_\mu \gamma_\mu$. Let's concentrate on the first term, collecting the factors of i, and using the

anti-commutation relation of the γ-matrices, $A\!\!\!/B\!\!\!/ = 2A.B - B\!\!\!/A\!\!\!/$, to write

$$i\bar{u}(p_1)g_s\epsilon\!\!\!/t^A\frac{(p\!\!\!/_1 + k\!\!\!/)}{(p_1 + k)^2}e_q\gamma_\mu v(p_2) = ig_s\bar{u}(p_1)\frac{[2\epsilon.(p_1 + k) - (p\!\!\!/_1 + k\!\!\!/)\epsilon\!\!\!/]}{(p_1 + k)^2}e_q\gamma_\mu t^A v(p_2)\,, \tag{13a}$$

$$\simeq ig_s\frac{p_1.\epsilon}{p_1.k}\,\bar{u}(p_1)e_q\gamma_\mu t^A v(p_2)\,, \tag{13b}$$

where to obtain the second line we have made use of the fact that $\bar{u}(p_1)p\!\!\!/_1 = 0$, $p_1^2 = k^2 = 0$, and taken the soft approximation $k_\mu \ll p_\mu$, which allows us to neglect the terms in the numerator that are proportional to k rather than p. The answer including both terms in Eq. (12) is

$$\mathcal{M}_{q\bar{q}g} \simeq \bar{u}(p_1)ie_q\gamma_\mu t^A v(p_2) \cdot g_s\left(\frac{p_1.\epsilon}{p_1.k} - \frac{p_2.\epsilon}{p_2.k}\right)\,, \tag{14}$$

where the first factor has the Lorentz structure of the $\mathcal{M}_{q\bar{q}}$ amplitude, i.e., apart from the colour matrix t^A, $\mathcal{M}_{q\bar{q}}$ is simply proportional to the $\mathcal{M}_{q\bar{q}}$ result. We actually need the squared amplitude, summed over polarizations and colour states,

$$
\begin{aligned}
|\mathcal{M}_{q\bar{q}g}|^2 &\simeq \sum_{A,a,b,\text{pol}}\left|\bar{u}_a(p_1)ie_q\gamma_\mu t^A v_b(p_2)\,g_s\left(\frac{p_1.\epsilon}{p_1.k} - \frac{p_2.\epsilon}{p_2.k}\right)\right|^2 \\
&= -|\mathcal{M}_{q\bar{q}}|^2 C_F g_s^2\left(\frac{p_1}{p_1.k} - \frac{p_2}{p_2.k}\right)^2 = |\mathcal{M}_{q\bar{q}}|^2 C_F g_s^2\frac{2p_1.p_2}{(p_1.k)(p_2.k)}\,.
\end{aligned} \tag{15}
$$

We have now explicitly written the quark colour indices a, b again. To obtain the second line we have made use of the result that $\sum_{A,a,b} t_{ab}^A t_{ba}^A = C_F N_C$ [cf. Eq. (8b)], whereas for $|\mathcal{M}_{q\bar{q}}|^2$ we have $\sum_{A,a,b}\delta_{ab}t_{ba}^A = N_C$. To carry out the sum over gluon polarizations we have exploited the fact that $\sum_{\text{pol}}\epsilon_\mu(k)\epsilon_\nu^*(k) = -g_{\mu\nu}$, plus terms proportional to k_μ and k_ν that disappear when dotted with the amplitude and its complex conjugate.

One main point of the result here is that in the soft limit, the $|\mathcal{M}_{q\bar{q}g}|^2$ squared matrix element *factorizes* into two terms: the $|\mathcal{M}_{q\bar{q}}|^2$ matrix element and a piece with a rather simple dependence on the gluon momentum.

The next ingredient that we need is the the phase space for the $q\bar{q}g$ system, $d\Phi_{q\bar{q}g}$. In the soft approximation, we can write this

$$d\Phi_{q\bar{q}g} \simeq d\Phi_{q\bar{q}}\frac{d^3\vec{k}}{2E(2\pi)^3}\,, \tag{16}$$

where $E \equiv E_k$ is the energy of the gluon k. We see that the phase space also factorizes. Thus we can write the full differential cross section for $q\bar{q}$ production plus soft gluon emission as the $q\bar{q}$ production matrix element and phase space, $|\mathcal{M}_{q\bar{q}}|^2 d\Phi_{q\bar{q}}$, multiplied by a soft gluon emission probability $d\mathcal{S}$,

$$|\mathcal{M}_{q\bar{q}g}|^2 d\Phi_{q\bar{q}g} \simeq |\mathcal{M}_{q\bar{q}}|^2 d\Phi_{q\bar{q}}d\mathcal{S}\,, \tag{17}$$

with

$$d\mathcal{S} = EdE\,d\cos\theta\,\frac{d\phi}{2\pi} \cdot \frac{2\alpha_s C_F}{\pi}\frac{2p_1.p_2}{(2p_1.k)(2p_2.k)}\,, \tag{18}$$

where we have used $d^3k = E^2 dE d\cos\theta d\phi$, expressing the result in terms of the polar (θ) and azimuthal (ϕ) angles of the gluon with respect to the quark (which itself is back-to-back with the antiquark, since we work in the centre-of-mass frame and there is negligible recoil from the soft gluon). With a little more algebra, we get our final result for the probability of soft gluon emission from the $q\bar{q}$ system

$$d\mathcal{S} = \frac{2\alpha_s C_F}{\pi}\frac{dE}{E}\frac{d\theta}{\sin\theta}\frac{d\phi}{2\pi}\,. \tag{19}$$

8

This result has two types of non-integrable divergence: one, called the soft (or infrared) divergence when $E \to 0$ and the other, a collinear divergence, when $\theta \to 0$ (or π), i.e., the gluon becomes collinear with the quark (or antiquark) direction. Though derived here in the specific context of $e^+ e^- \to q\bar{q}$ production, these soft and collinear divergences are a very general property of QCD and appear whenever a gluon is emitted from a quark, regardless of the process.

2.2   The total cross section

If we want to calculate the $\mathcal{O}(\alpha_s)$ corrections to the total cross section, the diagrams included in Eq. (12) are not sufficient. We also need to include a one-loop correction ('virtual'), specifically, the interference between one-loop $\gamma^* \to q\bar{q}$ diagrams and the tree-level $\gamma^* \to q\bar{q}$ amplitude, for example a contribution such as
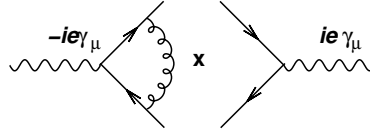
which has the same perturbative order (number of g_s factors) as the square of Eq. (12).

   The total cross section for the production of hadrons must be finite. The integral over the gluon-emission correction has two non-integrable, logarithmic divergences. These divergences must therefore somehow be cancelled by corresponding divergences in the virtual term. This is the requirement of *unitarity*, which is basically the statement that probability of anything happening must add up to 1. The most straightforward way of doing the full calculation for the total cross section is to use dimensional regularization in the phase space integral for the real emission diagram and for the integration over the loop momentum in the virtual diagram. However, in order just to visualize what is happening one can also write

$$\sigma_{tot} = \sigma_{q\bar{q}} \left(1 + \frac{2\alpha_s C_F}{\pi} \int \frac{dE}{E} \int \frac{d\theta}{\sin\theta} R\left(\frac{E}{Q}, \theta\right) - \frac{2\alpha_s C_F}{\pi} \int \frac{dE}{E} \int \frac{d\theta}{\sin\theta} V\left(\frac{E}{Q}, \theta\right) \right), \quad (20)$$

where the first term, 1, is the 'Born' term, i.e., the production of just $q\bar{q}$, the second term is the real emission term, and the third term is the loop correction. Since we need to integrate gluon emission beyond the soft and collinear region, we have introduced a function $R(E/Q, \theta)$, which parametrizes the deviation of the matrix-element from its soft limit when $E \sim Q$, with Q the centre-of-mass energy of the process. R has the property $\lim_{E \to 0} R(E/Q, \theta) = 1$. We have written the virtual term in a similar form, using $V(E/Q, \theta)$ to parametrize its structure (we cheat a bit, since the loop momentum integral includes regions of phase space where the gluon is offshell; this won't matter though for us here). The statement that real and virtual divergences cancel means that V should be identical to R in the soft or collinear limits

$$\lim_{E \to 0} (R - V) = 0, \qquad \lim_{\theta \to 0, \pi} (R - V) = 0. \quad (21)$$

Thus the corrections to the total cross section come from the region of hard ($E \sim Q$), large-angle gluons (for which perturbation theory is valid). There's a good reason for this: soft and collinear emission takes place on a time-scale $\sim 1/(E\theta^2)$ that is long compared to that, $\sim 1/Q$, for the production of the $q\bar{q}$ pair from the virtual photon. Anything that happens long after the production of the $q\bar{q}$ pair cannot change the fact that there will be a QCD final state (though it can change the properties of that final state), and so it does not affect the total cross section. Similarly, whatever dynamics is involved in effecting the transition between partons and hadrons is also expected to occur on a long time-scale ($\sim 1/\Lambda$) and so should not modify the total cross section. This is important because it allows us to neglect the issue that we cannot directly compute the properties of hadron production.

The fact that the corrections to the total cross section are dominated by a region of hard gluon emission is reflected in a reasonable behaviour for the perturbative series

$$\sigma_{tot} = \sigma_{q\bar{q}} \left(1 + 1.045 \frac{\alpha_s(Q)}{\pi} + 0.94 \left(\frac{\alpha_s(Q)}{\pi} \right)^2 - 15 \left(\frac{\alpha_s(Q)}{\pi} \right)^3 + \cdots \right) , \quad (22)$$

where we have expressed the result in terms of α_s evaluated at a renormalization $\mu = Q$ and the coefficients correspond to $Q = M_Z$ (for all known terms in the series, including electroweak effects, see Refs. [18–23] as well as references therein).

2.3 The final state

As a first step towards asking questions about the final state, our next exercise is to attempt to determine the mean number of gluons that are emitted from a quark with energy $\sim Q$. If the emission probability is small ($\propto \alpha_s$) then to first order in the coupling the mean number of emitted gluons is equal to the probability of emitting one gluon

$$\langle N_g \rangle \simeq \frac{2\alpha_s C_F}{\pi} \int^Q \frac{dE}{E} \int^{\pi/2} \Theta(E\theta > Q_0) . \quad (23)$$

The integral diverges for $E \to 0$ and $\theta \to 0$, however, we can reasonably argue that the divergent structure only holds as long as perturbation theory is valid. This motivates us to cut the divergences off at a scale $Q_0 \sim \Lambda$, because below that scale the language of quarks and gluons loses its meaning. That immediately tells us that we should have $E \gtrsim Q_0$, but it's not so immediately clear how the θ integral will be cut off. It turns out, for reasons to do with invariance of the small-angle emission pattern as one boosts the quark in the (longitudinal) direction of its motion, that the correct variable to cut on is *transverse momentum*, $k_t \sim E\theta$. We therefore find, to first order in the coupling,

$$\langle N_g \rangle \simeq \frac{\alpha_s C_F}{\pi} \ln^2 \frac{Q}{Q_0} + \mathcal{O}\left(\alpha_s \ln Q\right) , \quad (24)$$

where we have explicitly kept track only of the term with the largest number of logarithms. If we take $Q_0 = \Lambda$, how big is this result? We have to decide on the scale for α_s. Being hopelessly optimistic, i.e., taking $\alpha_s = \alpha_s(Q) = (2b \ln Q/\Lambda)^{-1}$ gives us
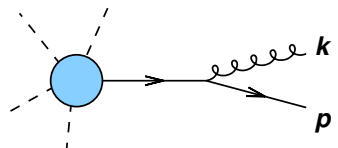
$$\langle N_g \rangle \simeq \frac{C_F}{2b\pi} \ln \frac{Q}{\Lambda} \simeq \frac{C_F}{4b^2 \pi \alpha_s} , \quad (25)$$

which, numerically, corresponds to $\langle N_g \rangle \simeq 2$. This is neither small numerically, nor parametrically ($\sim 1/\alpha_s$). Does this render perturbation completely useless for anything other than total cross sections?

We can follow two possible avenues to help answer this question. One approach is to calculate the next order, and see what structure it has. Alternatively we can ask whether there are final-state observables that have a better-behaved perturbative series than 'the mean number of gluons'.

2.3.1 Gluon (and hadron) multiplicity

Once one gluon has been emitted, it can itself emit further gluons. To understand what the gluon multiplicity might look like to higher orders, it's useful to write down the general pattern of emission of a soft gluon both from a quark and from a gluon, which is essentially independent of the process that produced the 'emitter':

$$\simeq \frac{2\alpha_s C_F}{\pi} \frac{dE}{E} \frac{d\theta}{\theta} , \quad (26a)$$
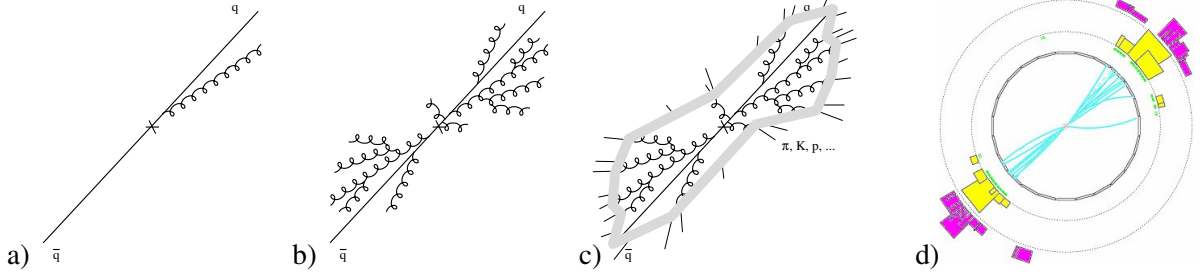
Fig. 5: Emission pattern from a $q\bar{q}$ event, with first a single gluon (a), then multiple emissions of gluons both from the $q\bar{q}$ pair and from the previously emitted gluons (b), followed by some process, 'hadronization', that causes hadrons to be produced from the gluons, giving an event (c), that structurally resembles a real event (d) ($e^+e^- \to Z \to$hadrons at LEP in the OPAL detector)

$$\simeq \quad \frac{2\alpha_s C_A}{\pi} \frac{dE}{E} \frac{d\theta}{\theta} . \tag{26b}$$

These expressions are valid when the emitted gluon is much lower in energy than the emitter, $k \ll p$, and when the emission angle θ is much smaller than the angle between the emitter and any other parton in the event (this is known as the condition of angular ordering [24]). The structure of emission of a soft gluon is almost identical from a quark and from a gluon, except for the substitution of the $C_F = 4/3$ colour factor in the quark case with the $C_A = 3$ colour factor in the gluon case.

Since quarks and gluons emit in similar ways, every gluon that is emitted from the quark can itself emit further gluons, and so forth. Most of the emissions will either be in almost the same direction as the original quark (due to the collinear divergence) and/or be soft. This is represented in Figs. 5(a) and (b) (for simplicity we've not shown gluons splitting to $q\bar{q}$ pairs, which also occurs, with just a collinear divergence). This still only gives a description of events in terms of quarks and gluons, whereas real events consist of hadrons. Though hadronization, the transition from quarks and gluon to hadrons is not something that we know how to calculate from first principles, one idea that has had some success is Local Parton Hadron Duality (LPHD) (see, e.g., Ref. [25]). It states that after accounting for all gluon and quark production down to scales $\sim \Lambda$, the transition from partons to hadrons is essentially local in phase space. Thus the hadron directions and momenta will be closely related to the partons', and the hadron multiplicity will reflect the parton multiplicity too. This is illustrated in Fig. 5(c), comparing it also to the picture of a real event, Fig. 5(d). The latter illustrates how the hadrons do tend to have the same collimated angular distribution as is predicted for gluons, with the small number of exceptions having low energy (i.e., soft) as can be seen from the larger curvature in the experiment's magnetic field.

This comparison with a single event is suggestive that our picture of gluon emission and hadronization might be reasonable. A more quantitative test can be obtained by calculating the number of emitted gluons. This requires the extension of Eqs. (23)–(25) to multiple gluon emission. The full calculation doesn't fit into the space available for these lectures (see instead textbook discussions in Refs. [1, 25]), but the basic idea is that there are terms $(\alpha_s \ln^2 Q/Q_0)^n$ for all orders n and that one can calculate their coefficients analytically. The structure of the result is

$$\langle N_g \rangle \sim \frac{C_F}{C_A} \sum_{n=1}^{\infty} \frac{1}{(n!)^2} \left(\frac{C_A}{2\pi b^2 \alpha_s} \right)^n \sim \frac{C_F}{C_A} \exp\left(\sqrt{\frac{2C_A}{\pi b^2 \alpha_s(Q)}} \right) , \tag{27}$$

where we've neglected to write the prefactor in front of the exponential, and we've also not given the subleading terms [26].
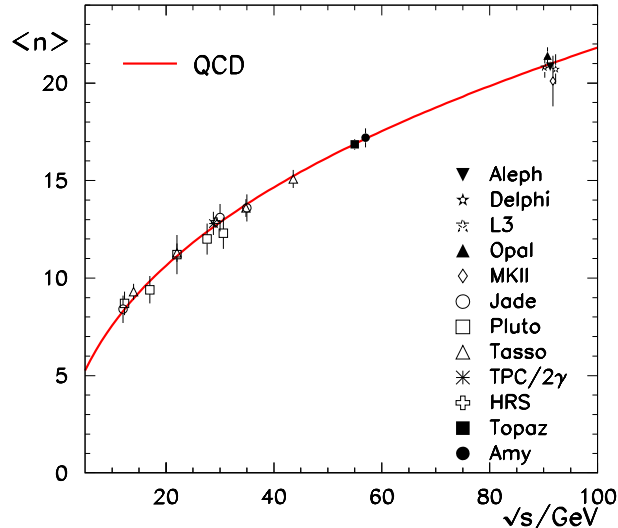
11

Fig. 6: Multiplicity of charged hadrons in $e^+e^- \rightarrow$ hadrons events, comparing the experimental data at a range of centre-of-mass energies Q, with the QCD prediction using a fitted normalisation and non-perturbative scale Λ. Figure adapted from Ref. [1].

How is Eq. (27) to be related to the hadron multiplicity? The simplest assumption is that each final parton gives some (unknown) fixed number of hadrons which must be fitted to data. Equation (27) then predicts not the total hadron multiplicity but its energy dependence. That prediction is illustrated in Fig. 6 and shows remarkable agreement with data over a range of energies, providing strong evidence that the picture outlined above is a fair reflection of 'reality'.

The above approach can be extended to calculate other properties of events such as the energy spectrum of hadrons, the fluctuations in the number of hadrons, and even correlations between hadrons, generally with reasonable success. However, as one probes more detailed features of events, the analytical calculations become significantly more complicated and one also becomes increasingly sensitive to the oversimplicity of the LPHD concept. Having said that, the same ideas that we are using, i.e., the importance of multiple soft and collinear splitting together with a transition from partons to hadrons, are, in numerical form, at the base of widely used Monte Carlo parton-shower event generators like PYTHIA, HERWIG and SHERPA. We will discuss them in more detail in Section 4.2.

2.3.2 *Infrared safe observables*

It is heartening that the above soft-collinear discussion gave such a good description of the data. However, it did involve the application of perturbation theory to kinematic regions where its validity is questionable, the need to calculate dominant contributions at all orders in $\alpha_{\rm s}$, and the introduction of a free parameter to 'fudge' the fact that we don't understand the non-perturbative physics. A natural question is therefore whether one can formulate final-state observables for which these problems are not present.

The answer is that one can. For an observable to be calculated based on just one or two orders of perturbation theory it should be infrared and collinear (IRC) safe. In the words of Ref. [1]:

> For an observable's distribution to be calculable in [fixed-order] perturbation theory, the observable should be infra-red safe, i.e. insensitive to the emission of soft or collinear gluons. In particular if $\vec{p}_i$ is any momentum occurring in its definition, it must be invariant under the branching
>
> $$\vec{p}_i \rightarrow \vec{p}_j + \vec{p}_k$$
>
> whenever $\vec{p}_j$ and $\vec{p}_k$ are parallel [collinear] or one of them is small [infrared].

12

For example, the multiplicity of gluons is not IRC safe, because it is modified by soft and collinear splitting. The energy of the hardest particle in an event is not IRC safe, because it is modified by collinear splitting. However, the total energy flowing into a given cone is IRC safe, because soft emissions don't modify the energy flow, and collinear emissions don't modify its direction.

This last example comes from Sterman and Weinberg [27], who defined an e^+e^- event as having 2 'jets' if at least a fraction $(1 - \epsilon)$ of the event's energy is contained in two cones of half-angle δ. If we take δ to be $30°$ and $\epsilon = 0.1$, then Fig. 5(d) is an example of such a 2-jet event. We can adapt our expression for the total cross section, Eq. (22), to give us the 2-jet cross section as follows

$$\sigma_{\text{2-jet}} = \sigma_{q\bar{q}} \left(1 + \frac{2\alpha_{\text{s}} C_F}{\pi} \int \frac{dE}{E} \int \frac{d\theta}{\sin \theta} \left[R\left(\frac{E}{Q}, \theta\right) \left(1 - \Theta\left(\frac{E}{Q} - \epsilon\right) \Theta(\theta - \delta)\right) - V\left(\frac{E}{Q}, \theta\right) \right] \right). \tag{28}$$

For small E or θ this is just like the total cross section, with full cancellation of divergences between real and virtual terms [cf. Eq. (21)]. For large E and large θ a *finite* piece of real-emission cross section is cut out by the factor $(1 - \Theta(\frac{E}{Q} - \epsilon)\Theta(\theta - \delta))$ and it corresponds to scales with $E \sim Q$ and large angles, for which perturbation theory is valid. This then gives

$$\sigma_{\text{2-jet}} = \sigma_{q\bar{q}}(1 - c_1\alpha_{\text{s}} + c_2\alpha_{\text{s}}^2 + \cdots), \tag{29}$$

where c_1, c_2, etc. are all of order 1 (as long as ϵ and δ were not taken too small). Similarly one could define a 3-jet cross section by requiring that it not be a 2-jet event and that all but a fraction ϵ of the energy be contained in 3 cones of half angle δ. This would give a cross section of the form

$$\sigma_{\text{3-jet}} = \sigma_{q\bar{q}}(c_1'\alpha_{\text{s}} + c_2'\alpha_{\text{s}}^2 + \cdots), \tag{30}$$

where, again, the coefficients are all $\mathcal{O}(1)$. So whereas the cross section for getting an extra gluon is divergent, the cross section for an extra jet is finite and small, $\mathcal{O}(\alpha_{\text{s}})$. One difficulty with the extension of the Sterman–Weinberg definition to multiple jets is to know how to place the cones. Since jet-finding is a well-developed subject in its own right, we will return to in detail in Section 5.

The Sterman–Weinberg jet cross section gives a discrete classification of events: an event either has two jets, or more. An alternative class of infrared and collinear safe observables is that of event shapes, which give a continuous classification of events. The most widely studied example is the thrust, T,

$$T = \max_{\vec{n}_T} \frac{\sum_i |\vec{p}_i.\vec{n}_T|}{\sum_i |\vec{p}_i|}, \tag{31}$$

where the sum runs over all particles, and one chooses the thrust axis $\vec{n}_T$ (a 3-dimensional unit vector) so as to maximize the projection in the numerator. For a perfectly collimated 2-jet event, the thrust axis aligns with the jet axes and the thrust is 1. For a 'Mercedes' type event with three identical collimated jets, the thrust axis will be aligned with any one of the three jets and the thrust will be $2/3$. Intermediate events will have intermediate values of the thrust.

One application of the thrust variable is given in Fig. 7. It shows data for the thrust distribution from LEP, compared to $\mathcal{O}(\alpha_{\text{s}})$ calculations of the thrust distribution in QCD and in a variant of QCD in which the gluon is a scalar particle rather than a vector particle. The scalar gluon case does not have a divergence for soft emission (the collinear divergence is still there), with the result is that the distribution diverges less strongly in the 2-jet limit than for vector gluons. The data clearly prefer the vector-gluon case, though they do also show the need for higher-order corrections at thrust values close to $2/3$ and 1.

More generally, event shapes like the thrust have seen broad use in measurements of the QCD coupling, tuning of Monte Carlo event generators (see Section 4.2), studies of the hadronization process, and also as an event-topology discriminant in searches for decays of particles beyond the Standard Model.
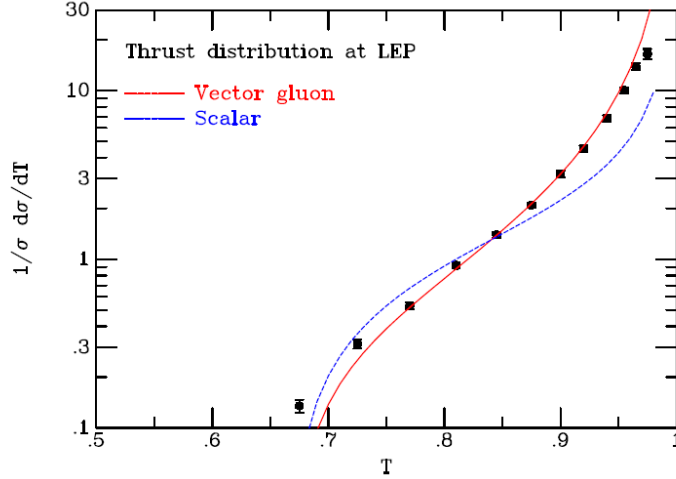
Fig. 7: Measured thrust distribution at LEP compared to leading order predictions based on QCD (vector gluon, solid red line) and a modified version of QCD in which the gluon is a scalar (spin 0, dashed blue line) rather than a vector (spin-1) particle. Figure taken from CERN academic training lectures by B. R. Webber.

2.4 Summary

The $e^+e^- \to$ hadrons process has allowed us to examine many of the basic ideas of perturbative QCD: soft and collinear divergences, the question of which observables are perturbatively calculable or not (related to infrared and collinear safety) and even what happens if one takes perturbation theory seriously outside its strict domain of applicability (one acquires a rough understanding of the collimated, high-multiplicity structure of real events).

3 Parton distribution functions

Having considered processes that involve hadrons in the final state, let us now examine what happens when they are present in the initial state. The importance of understanding initial-state hadrons is obvious at the LHC. Within the 'parton model', we write, for example, the hadron collider cross section to produce a Z and a Higgs boson as

$$\sigma = \int dx_1 f_{q/p}(x_1) \int dx_2 f_{\bar{q}/p}(x_2) \, \hat{\sigma}_{q\bar{q} \to ZH}(x_1 x_2 s) \,, \tag{32}$$

where $s = (p_1 + p_2)^2$ is the squared pp centre-of-mass energy, $f_{q/p}(x_1)$ is the number density of quarks of type q carrying a fraction x_1 of the momentum of the first proton, and similarly with $f_{\bar{q}/p}(x_2)$ for the other proton. The $f_{q/p}(x)$ functions are known as 'parton distribution functions' (PDFs). They multiply the 'hard' (here, electroweak) cross section, $\hat{\sigma}_{q\bar{q} \to ZH}(x_1 x_2 s)$ for the process $q\bar{q} \to ZH$, a function of the squared partonic ($q\bar{q}$) centre-of-mass energy, $\hat{s} = x_1 x_2 s$. After integrating over x_1 and x_2 (and summing over quark species), one obtains the total cross section for $pp \to ZH$. The above form seems quite intuitive, but still leaves a number of questions: for example, how do we determine the momentum distributions of quarks and gluons inside the proton? How does the 'factorization' into PDFs and a hard part stand up to QCD corrections? Understanding these questions is crucial if, one day, we are to take measured cross sections for ZH and interpret them, for example, in terms of the coupling of the Higgs to the Z. And they're just as crucial for practically any other physics analysis we might want to carry out at the LHC.
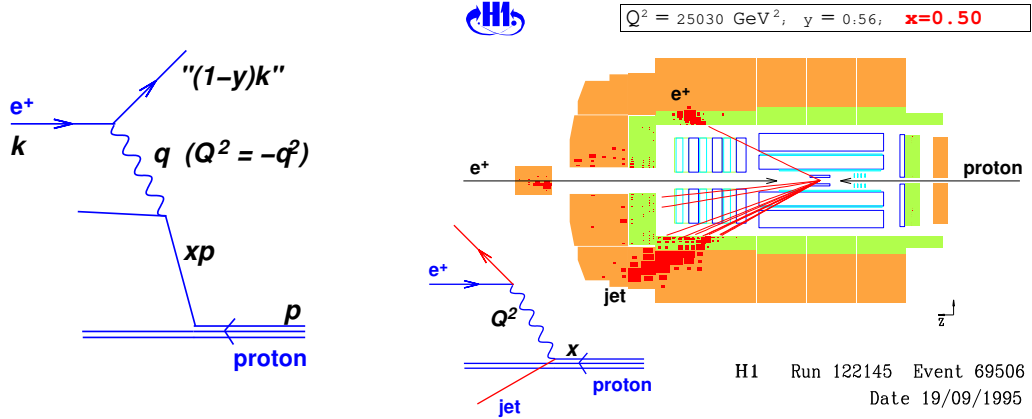
14

Fig. 8: Left: kinematic variables of DIS; right: illustration of an event as it appeared in practice in the H1 detector at HERA.

The parton distribution functions are properties of the (non-perturbative) proton. A natural first question is whether we can calculate the PDFs based on lattice QCD. In principle, yes, and there is ongoing work in this direction (see, e.g., Ref. [28]), however, currently lattice QCD has not reached an accuracy in these calculations that is competitive with the combination of experimental measurements and perturbative QCD analyses that we discuss below.

3.1 Deep Inelastic Scattering

The process where we have learnt the most about PDFs is Deep Inelastic Scattering (DIS), i.e., lepton–proton scattering in which the photon that is exchanged between lepton and proton has a large virtuality. The kinematics of the "quark-parton-model" DIS process is represented in Fig. 8 (left) and an event from the H1 detector at HERA is shown on the right. Kinematic variables that are usually defined are

$$Q^2 = -q^2\,, \qquad x = \frac{Q^2}{2p.q}\,, \qquad y = \frac{p.q}{p.k}\,, \tag{33}$$

where Q^2 is the photon virtuality, x is the longitudinal momentum fraction of the struck quark in the proton, and y is the momentum fraction lost by the electron (in the proton rest frame).

To zeroth order in α_s (the 'quark parton model'), the DIS cross section can be written as

$$\frac{d^2\sigma^{em}}{dxdQ^2} = \frac{4\pi\alpha^2}{xQ^4}\left(\frac{1+(1-y)^2}{2}F_2^{em} + \mathcal{O}\left(\alpha_\mathrm{s}\right)\right)\,, \tag{34}$$

written in terms of the F_2 structure function, which, to zeroth order in α_s is given by

$$F_2^{em} = x\sum_{i=q,\bar{q}} e_i^2 f_{i/p}(x) + \mathcal{O}\left(\alpha_\mathrm{s}\right)\,, \tag{35}$$

(the "em" superscript is a reminder that we're only considering the electromagnetic part here).

Given the sum over all flavours in Eq. (35), disentangling the information about individual flavours might seem like quite a daunting task. We can attempt to see where the information comes from by starting off with the assumption that the proton consists just of up and down quarks, in which case

$$F_2^{\mathrm{proton}} = x(e_u^2 u_p(x) + e_d^2 d_p(x)) = x\left(\frac{4}{9}u_p(x) + \frac{1}{9}d_p(x)\right)\,, \tag{36}$$

where we have introduced the shorthand $f_{u/p}(x) = u_p(x)$, etc. (later we will drop the "p" subscript altogether). In Eq. (36) we now have a linear combination of just two functions. The next step is to use
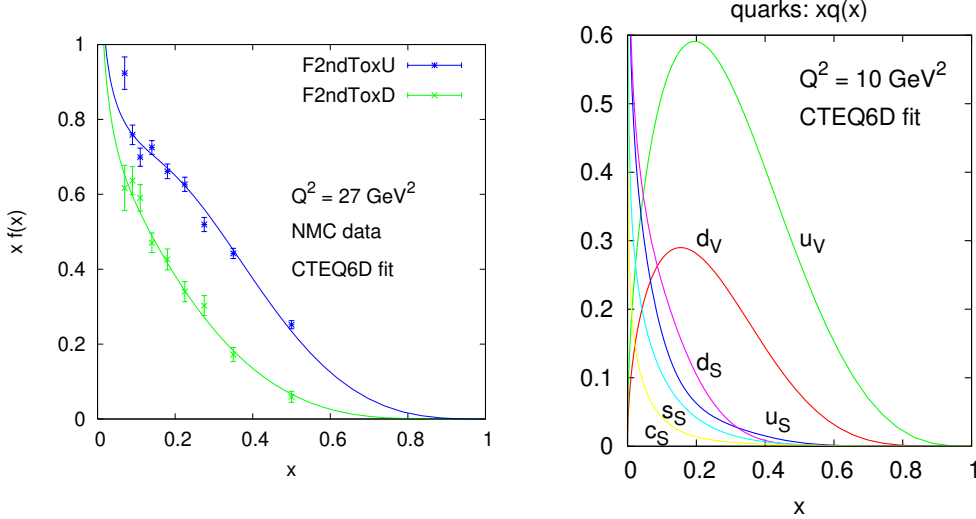
15

Fig. 9: Left: linear combinations of NMC F_2 data [29] for protons and deuterons so as to obtain $xu(x)$ and $xd(x)$, assuming only u and d quarks inside the proton, together with the expectations from the same linear combination based on the CTEQ6D PDF parametrizations [30]. Right: results for different valence and sea quark distributions from CTEQ6D at $Q^2 = 10\,\mathrm{GeV}^2$.

isospin symmetry, the fact that the neutron is essentially just a proton with $u \leftrightarrow d$, i.e., $u_n(x) \simeq d_p(x)$ (ignoring small electromagnetic effects), so that

$$\frac{1}{x}F_2^{\text{neutron}} = \frac{4}{9}u_n(x) + \frac{1}{9}u_n(x) \simeq \frac{4}{9}d_p(x) + \frac{1}{9}u_p(x)\,. \tag{37}$$

Appropriate linear combination of F_2^{proton} and F_2^{neutron} (in practice one uses deuterons, or nuclei as a source of neutrons) therefore provides separate information on $u_p(x)$ and $d_p(x)$.

The results of this exercise are shown in Fig. 9 (left). As expected we see more up quarks than down quarks. But there's also a problem: let's try to extract the total number of up quarks, $U = \int_0^1 dx\, u(x)$. We see that the data increase towards small x, and a parametrization (CTEQ6D [30]) of these and other data even seems to diverge as $x \to 0$, $xu(x) \sim xd(x) \sim x^{-0.25}$.[4] Given that the plot is supposed to be for $xu(x)$ and that we need to integrate $u(x)$ to get the total number of up quarks in the proton, it looks like we'll obtain an infinite number of up quarks, which is hardly consistent with expectations from the picture of a proton as being a uud state.

One thing we've 'neglected' is that there can also be anti-up and anti-down quarks in the proton, because the proton wavefunction can fluctuate, creating $u\bar{u}$ and $d\bar{d}$ pairs, 'sea quarks', and so give rise to $\bar{u}(x)$ and $\bar{d}(x)$ distributions. Therefore instead of Eq. (36), we should have written

$$F_2^{\text{proton}} = \frac{4}{9}(xu_p(x) + x\bar{u}_p(x)) + \frac{1}{9}(d_p(x) + \bar{d}_p(x))\,, \tag{38}$$

since quarks and antiquarks have identical squared charges. So what we called "$xu(x)$" in Fig. 9 (left) was actually $xu(x) + x\bar{u}(x)$ (with some admixture of strange and charm quarks too). The infinite number of quarks and antiquarks can then just be interpreted as saying that fluctuations within the proton create infinite numbers of $q\bar{q}$ pairs, mostly carrying a small fraction of the proton's momentum.

Returning to the statement that the proton has 2 up and 1 down quark, what we mean is that the

---

[4]The particular parametrization shown here is not very widely used nowadays, but is particularly convenient for comparisons to DIS data (it corresponds to a definition of PDFs known as the "DIS scheme," for which higher-order corrections to many DIS cross sections are particularly simple). The behaviour at small x is common to essentially all parametrizations.

net number of up minus anti-up quarks is 2,

$$\int_0^1 dx(u(x) - \bar{u}(x)) = 2\,, \qquad \int_0^1 dx(d(x) - \bar{d}(x)) = 1\,, \tag{39}$$

where $u(x) - \bar{u}(x)$ is also called the valence quark distribution $u_V(x)$. How can we measure the difference between quarks and antiquarks? The answer is through charged-current processes (e.g., neutrino scattering), since a W^+ interacts only with d and $\bar{u}$, but not with $\bar{d}$ or u.

We could imagine forming linear combinations of data sets with proton and nuclear targets, with photon and $W^\pm$ exchange, etc., in order to obtain the different quark-flavour PDFs. In practice it is simpler to introduce parametrizations for each flavour, deduce predictions for DIS, neutrino-scattering and other cross sections, and then fit the parameters so as to obtain agreement with the experimental data. This is known as a 'global fit'. We will return to such fits below, but for now let's just look at the results for different quark distributions, shown in Fig. 9 (right).

We see that the valence quark distributions are mainly situated at moderate x values. For $x \to 1$ they fall off as a moderate power of $(1 - x)$, roughly as $(1 - x)^3$. This fall-off comes about because to have one u quark carrying most of the proton momentum implies that the other u and the d quark must both carry only a small fraction, and the phase space for that to occur vanishes as $x \to 1$ (the proper treatment for this is through 'quark-counting rules'). For $x \to 0$, $xq_V(x) \sim x^{0.5}$, which is related to something called Regge physics.

The sea-quark distributions ($u_S(x) \equiv u(x) + \bar{u}(x) - u_V(x) \equiv 2\bar{u}(x)$, etc.) are concentrated at small x values. They fall off much more steeply at large x: for a $\bar{u}$ quark to have a large momentum fraction, 3 u-quarks and one d-quark must have correspondingly small-x, so the phase space is even more restricted than in the case of valence quarks. For $x \to 0$, the exact behaviour is not entirely simple, but as a rough rule of thumb $xq_S(x) \sim x^{-\omega}$ with $\omega \sim 0.2$–0.4.

Of course we're still missing the PDFs for one kind of parton: gluons. It's a well known fact that if we evaluate the fraction of the proton's momentum carried by all the quarks,
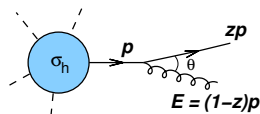
$$\sum_q \int_0^1 dx\, xq(x)\,, \tag{40}$$

the result is about 0.5. It's fair to suspect that the gluon is responsible for the other half, but how are we to establish its shape given that it's not directly probed by photons or $W^\pm$? To answer that question we must go beyond the 'naive' leading-order partonic picture of the proton's quarks interacting with probes, and bring in QCD splitting.

3.2 Initial-state parton splitting, DGLAP evolution

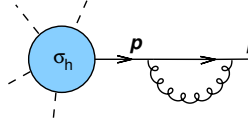### 3.2.1 Final and initial-state divergences

In Eq. (26a) we wrote the universal form for the final-state 'splitting' of a quark into a quark and a soft gluon. Let's rewrite it with different kinematic variables, considering a hard process h with cross section σ_h, and examining the cross section for h with an extra gluon in the final state, σ_{h+g}. We have

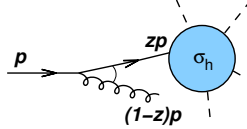$$\sigma_{h+g} \simeq \sigma_h \frac{\alpha_s C_F}{\pi} \frac{dz}{1-z} \frac{dk_t^2}{k_t^2}\,, \tag{41}$$

where E in Eq. (26a) corresponds to $E = (1-z)p$ and we've introduced $k_t = E\sin\theta \simeq E\theta$. If we avoid distinguishing a collinear $q + g$ pair from a plain quark (measurements with IRC safe observables) then, as we argued before, the divergent part of the gluon emission contribution always cancels with a
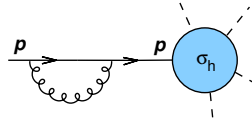
related virtual correction

$$\sigma_{h+V} \simeq -\sigma_h \frac{\alpha_s C_F}{\pi} \frac{dz}{1-z} \frac{dk_t^2}{k_t^2} . \qquad (42)$$

Now let us examine what happens for initial-state splitting, where the hard process occurs *after* the splitting and the momentum entering the hard process is modified $p \to zp$:

$$\sigma_{g+h}(p) \simeq \sigma_h(zp) \frac{\alpha_s C_F}{\pi} \frac{dz}{1-z} \frac{dk_t^2}{k_t^2} , \qquad (43)$$

where we have made explicit the hard process's dependence on the incoming momentum, and we assume that σ_h involves momentum transfers $\sim Q \gg k_t$, so that we can ignore the extra transverse momentum entering σ_h. For virtual terms, the momentum entering the process is unchanged, so we have
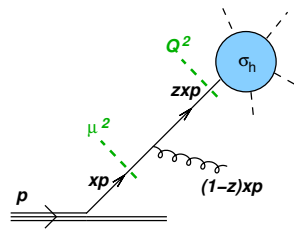

$$\sigma_{g+h}(p) \simeq -\sigma_h(p) \frac{\alpha_s C_F}{\pi} \frac{dz}{1-z} \frac{dk_t^2}{k_t^2} , \qquad (44)$$

The total cross section then gets contributions with two different hard cross sections:

$$\sigma_{g+h} + \sigma_{V+h} \simeq \frac{\alpha_s C_F}{\pi} \underbrace{\int_0^{Q^2} \frac{dk_t^2}{k_t^2}}_{\text{infinite}} \underbrace{\int_0^1 \frac{dz}{1-z} [\sigma_h(zp) - \sigma_h(p)]}_{\text{finite}} . \qquad (45)$$

Note the limits on the integrals, in particular the Q^2 upper limit on the transverse-momentum integration: the approximations we're using are valid as long as the transverse momentum emitted in the initial state is much smaller than the momentum transfers Q that are present in the hard process. Of the two integrations in Eq. (45), the one over z is finite, because in the region of the soft divergence, $z \to 1$, the difference of hard cross sections, $[\sigma_h(zp) - \sigma_h(p)]$, tends to zero. In contrast, the k_t integral diverges in the collinear limit: the cross section with an incoming parton (and virtual corrections) appears not to be collinear safe. This is a general feature of processes with incoming partons: so how are we then to carry out calculations with initial-state hadrons?

In Section 2.3.1, when trying to make sense of final-state divergences, we introduced a (non-perturbative) cutoff. Let's do something similar here, with a cutoff, μ_F, called a factorization scale (which will usually be taken at perturbative scales). The main idea in using this cutoff is that any emissions that occur with $k_t \lesssim \mu_F$ are absorbed ('factorized') into the PDF itself. Thus the PDFs become a function of μ_F. We can write the lowest order term for the cross section, σ_0, and the correction with one initial-state emission or virtual loop, σ_1, as follows

$$\sigma_0 = \int dx \, \sigma_h(xp) \, q(x, \mu_F^2) , \qquad (46a)$$

$$\sigma_1 \simeq \frac{\alpha_s C_F}{\pi} \underbrace{\int_{\mu_F^2}^{Q^2} \frac{dk_t^2}{k_t^2}}_{\text{finite (large?)}} \underbrace{\int \frac{dx \, dz}{1-z} [\sigma_h(zxp) - \sigma_h(xp)] \, q(x, \mu_F^2)}_{\text{finite}} , \qquad$$

$$(46b)$$

where we have now included also the integral over the longitudinal momentum fraction x of the parton extracted from the proton. The emissions and virtual corrections with $k_t \lesssim \mu_F$ are now implicitly included inside the $q(x, \mu_F^2)$ PDF factor that appears in the σ_0 contribution, and only those with $k_t \gtrsim \mu_F$

arise explicitly in the $\mathcal{O}(\alpha_\mathrm{s})$ term. This term (whose real-emission part is represented in the diagram to the left) is now finite, albeit potentially large if $\mu_\mathrm{F} \ll Q$.

This situation of having a non-integrable divergence that somehow needs to be regularized and absorbed with a scale choice into some 'constant' of the theory (here the PDFs), is reminiscent of renormalization for the coupling constant. The differences are that here we are faced with a divergence in the (collinear) infrared rather than one in the ultraviolet. And that unlike the coupling, the PDFs are not fundamental parameters of the theory, but rather quantities that we could calculate if only we had sufficiently sophisticated theoretical 'technology'. Nevertheless, as for the coupling, the freedom in choosing the scale that enters the regularization, called the factorization scale, implies the presence of a renormalization group equation for the PDFs, the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) equation.

3.2.2 The DGLAP equation

To see what form the DGLAP equation takes, let us fix the longitudinal momentum of the quark entering the hard process to be xp (whereas above we'd fixed the momentum for the quark extracted from the proton). Next we examine the effect on the PDFs of integrating over a small region of transverse momentum $\mu_\mathrm{F}^2 < k_t^2 < (1+\epsilon)\mu_\mathrm{F}^2$,

$$\frac{dq(x,\mu_\mathrm{F}^2)}{d\ln\mu_\mathrm{F}^2} = \frac{1}{\epsilon} \left(\quad \right) \tag{47}$$

$$= \frac{\alpha_\mathrm{s}}{2\pi} \int_x^1 dz\, p_{qq}(z)\, \frac{q(x/z,\mu_\mathrm{F}^2)}{z} \;-\; \frac{\alpha_\mathrm{s}}{2\pi} \int_0^1 dz\, p_{qq}(z)\, q(x,\mu_\mathrm{F}^2)$$

where $p_{qq}(z)$ is the real part of the 'splitting kernel' for a quark splitting to a quark plus a gluon,

$$p_{qq}(z) = C_F \frac{1+z^2}{1-z} \,. \tag{48}$$

Until now, we had concentrated on the soft limit, which was equivalent to approximating $p_{qq}(z) \simeq \frac{2C_F}{1-z}$. What Eq. (47) tells us is that as we increase the factorization scale, we get extra partons with longitudinal momentum fraction x that come from the branching of partons in the proton at lower factorization scales but larger momentum fractions x/z ($x < z < 1$). There are also loop contributions (second term on the RHS) to the parton density at a fixed x value, which are negative contributions to the evolution. The way to think about these is that when a parton with momentum fraction x branches to partons with lower momentum fractions, the original parton is lost and the loop diagram accounts for that.

It's a bit awkward to write the real and virtual parts separately in Eq. (47), especially if one wants to explicitly see the cancellation of the divergences for $z \to 1$. It's therefore standard to use the more compact notation

$$\frac{dq(x,\mu_\mathrm{F}^2)}{d\ln\mu_\mathrm{F}^2} = \frac{\alpha_\mathrm{s}}{2\pi} \underbrace{\int_x^1 dz\, P_{qq}(z)\, \frac{q(x/z,\mu_\mathrm{F}^2)}{z}}_{P_{qq}\otimes q}, \qquad P_{qq} = C_F \left(\frac{1+z^2}{1-z} \right)_+ , \tag{49}$$

where the subscript plus, known as the 'plus' prescription, is to be understood as follows:

$$\int_x^1 dz\, [g(z)]_+\, f(z) = \int_x^1 dz\, g(z)\, f(z) - \int_0^1 dz\, g(z)\, f(1) \tag{50}$$

$$= \int_x^1 dz\, g(z)\, (f(z) - f(1)) - \int_0^x dz\, g(z)\, f(1) \tag{51}$$
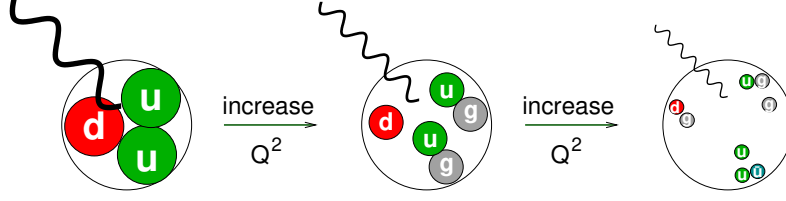
19

Fig. 10: An illustration how with ever shorter wavelength photon probes, one resolves more and more structure inside the proton.

so that the factor $(f(z) - f(1))$, which goes to zero at $z = 1$, kills the divergence due the singular behaviour of $g(z)$ for $z \to 1$.

Equation (49) involves just quarks, but the proton contains both quarks and gluons, so the full DGLAP equations are actually coupled evolution equations. Schematically, for just a single quark flavour, they read

$$\frac{d}{d \ln \mu_F^2} \begin{pmatrix} q \\ g \end{pmatrix} = \frac{\alpha_s(\mu_F^2)}{2\pi} \begin{pmatrix} P_{q \leftarrow q} & P_{q \leftarrow g} \\ P_{g \leftarrow q} & P_{g \leftarrow g} \end{pmatrix} \otimes \begin{pmatrix} q \\ g \end{pmatrix} \qquad (52)$$

and more generally they span all quark flavours and anti-flavours. In labelling the different flavour entries, we've included arrows (usually not shown), e.g. $q \leftarrow g$, so as to emphasize that we have evolution from the right-hand parton type to the left-hand parton type. The splitting functions other than P_{qq} are given by

$$P_{qg}(z) = T_R \left[z^2 + (1-z)^2 \right], \qquad P_{gq}(z) = C_F \left[\frac{1 + (1-z)^2}{z} \right], \qquad (53a)$$

$$P_{gg}(z) = 2C_A \left[\frac{z}{(1-z)_+} + \frac{1-z}{z} + z(1-z) \right] + \delta(1-z) \frac{(11C_A - 4n_f T_R)}{6} . \qquad (53b)$$

Additionally, $P_{\bar{q}g} = P_{qg}$ and, to this first order in the coupling, $P_{qq'}$ and $P_{q\bar{q}}$ are both zero.

Several features of the splitting functions are worth noting: P_{qg} and P_{gg} are both symmetric in $z \leftrightarrow 1 - z$ (except for the virtual part). P_{qq} and P_{gg} diverge for $z \to 1$, which corresponds to soft-gluon emission. And P_{gg} and P_{gq} both diverge for $z \to 0$ (corresponding to a soft gluon entering the 'hard' process). This last point implies that PDFs $q(x)$ and $g(x)$ must grow at least as fast as $1/x$ for $x \to 0$: even if such a divergence is absent in the non-perturbative 'initial conditions' for the quark and gluon distributions at low scales μ_F, DGLAP evolution inevitably introduces it into the result for $q(x, \mu_F^2)$ and $g(x, \mu_F^2)$ at higher scales.

3.2.3 Results of DGLAP evolution

Pictorially, the effect of DGLAP evolution is illustrated in Fig. 10. A more quantitative view is given in Fig. 11, which shows the effect of DGLAP evolution with an initial condition that is pure quark (two left plots) or pure gluon (two right plots). In both cases one sees that evolution generates some amount of the missing parton type; one also sees how it depletes the parton distributions at large x, and increases them at small x (especially in the case of the gluon). The attentive reader may have observed that the figure labels the scale of the PDFs as Q^2 rather than μ_F^2: this is because it is standard to take $\mu_F^2 = Q^2$ so as to minimize the size of the the $\mathcal{O}(\alpha_s)$ term of Eq. (46) which arises, roughly, from the integral over transverse momenta from μ_F^2 to Q^2. I.e., one usually chooses to factorize essentially *all* initial-state radiation into the PDFs and so into the LO cross section.

Since, as we've see in Fig. 11, the presence of a gluon distribution helps drive quark evolution, we can use the experimentally observed pattern of quark evolution to help constrain the gluon. The left-hand plot of Fig. 12 shows data from ZEUS [31] and NMC [29] on $F_2(x, Q^2)$ at some low but still perturbative
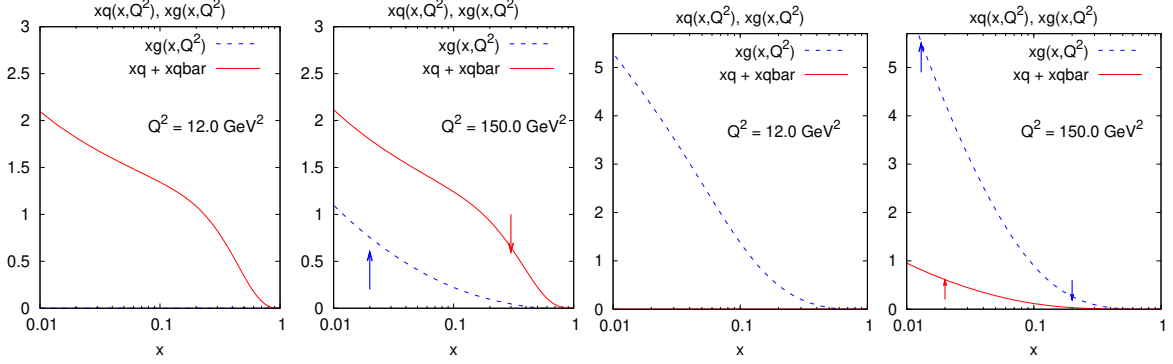
Fig. 11: An illustration of the impact of DGLAP evolution. From left to right: (a) initial condition consisting just of quarks and anti-quarks at $\mu_F^2 \equiv Q^2 = 12\,\mathrm{GeV}^2$; (b) the result of evolution to $Q^2 = 150\,\mathrm{GeV}^2$; (c) a purely gluonic initial condition at $Q^2 = 12\,\mathrm{GeV}^2$; and (d) the result of its evolution to $Q^2 = 150\,\mathrm{GeV}^2$.
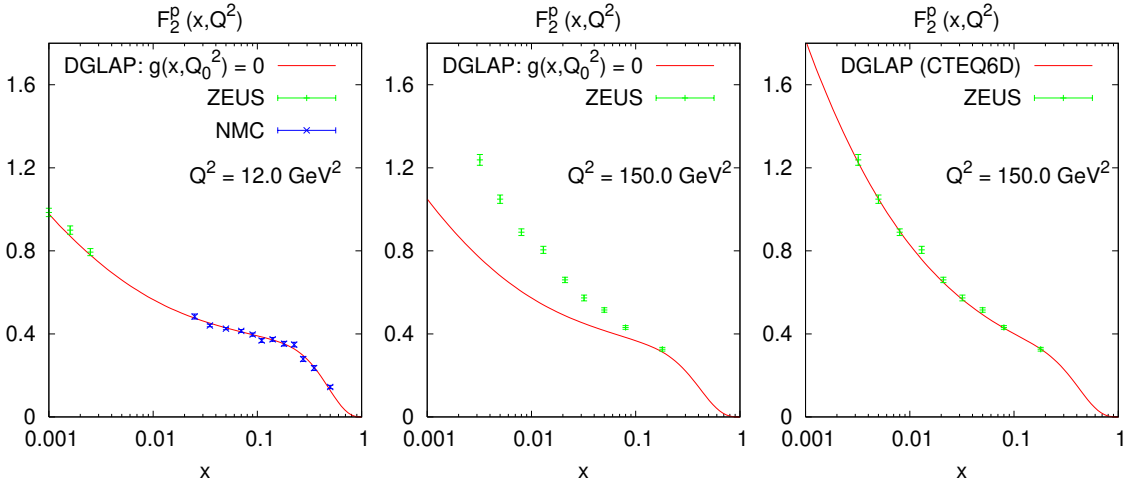

Fig. 12: ZEUS and NMC data together with an initial condition that gives a good fit at low Q^2 (left-most plot). If one evolves that initial condition assuming a gluon distribution that is zero at low Q^2, then agreement with high-scale data is poor (central plot); whereas with a significant low-scale gluon component (taken from the CTEQ6D parametrization), agreement becomes good at high scales (right-most plot).

scale $Q^2 = Q_0^2 \equiv 12\,\mathrm{GeV}^2$. The data are compared to the expectations based on the CTEQ6D PDFs' quark content at that scale, illustrating the good agreement. Since these are data for F_2, they have no direct sensitivity to the gluon distribution. The middle plot shows data for $150\,\mathrm{GeV}^2$, together with the results of DGLAP evolution from $Q_0^2 = 12\,\mathrm{GeV}^2$, assuming that the gluon distribution was zero at Q_0^2. There's a clear discrepancy. In the right-hand plot, the comparison is made with evolution whose initial condition at Q_0^2 contained a quite large gluon component (exactly that in the CTEQ6D distributions), causing the quark distribution at small x values to increase faster with Q^2 than would otherwise be the case, bringing the evolution into agreement with the data.

3.3 Global fits

It's interesting to ask just how much of a gluon distribution is needed in order to get the agreement shown in Fig. 12. The answer is given in Fig. 13 and one sees that the gluon distribution is *enormous*, especially at small values of x. It is fair to ask whether we can trust a result such as Fig. 13, so in this section we will examine some of ingredients and issues that are relevant to the 'global fits' that inform our knowledge of PDFs.

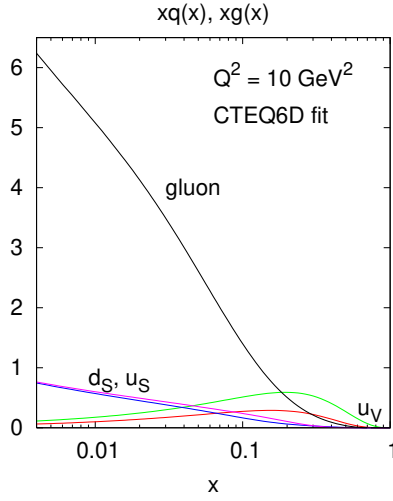**Fig. 13:** The distributions of different parton species in the CTEQ6D parametrization at a scale $Q^2 = 10\,\mathrm{GeV}^2$
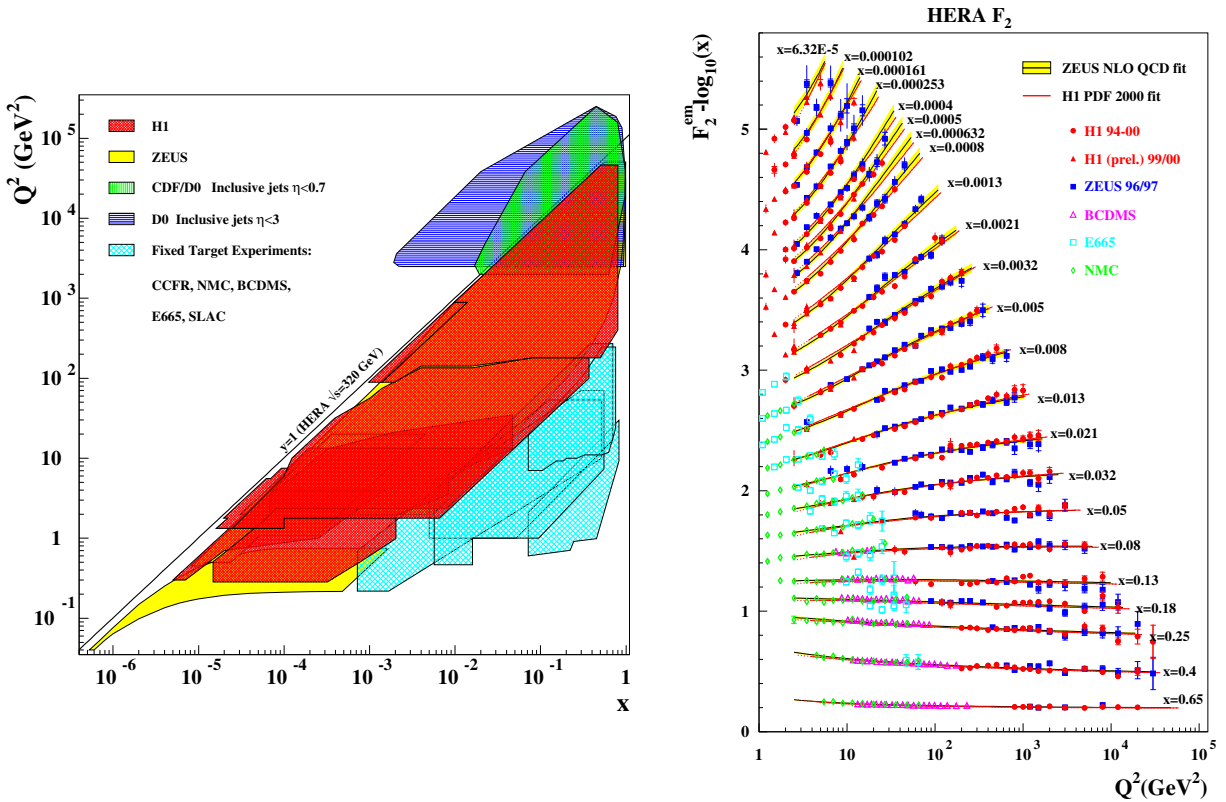

Fig. 14: Left: an illustration of the kinematic regions and data sets typically used in PDF fits (based on a fit by ZEUS). Right: experimental results for F_2 as a function of Q^2 for many different x values, compared to the results of a global fit by the ZEUS collaboration.

Figure 14 (left) illustrates the kinematical regions in the x and Q^2 plane covered by the experimental data sets typically used in global fits. Everything below the diagonal line corresponds to DIS data, and the right-hand plot shows the comparison between a fit (by ZEUS) and the bulk of the DIS data, illustrating the excellent consistency between fit and data. Agreement with such a broad data set is already a non-trivial achievement. Figure 14 (left) also shows shaded regions that span the diagonal line. These correspond to hadron-collider jets data, which provide valuable direct information on the gluon distribution in global fits, as we will discuss below. The other topics that we will address here relate to the accuracy of our determinations of PDFs.

3.3.1 Factorization and $p\bar{p}$ jet production

Perhaps the most convincing cross-check of PDF extractions comes from Tevatron jet data (there are also important jet data from HERA). The process of factorizing initial-state radiation into the PDF at a given scale is equally valid in DIS and $p\bar{p}$ (or pp) collisions, as illustrated in the following pictorial representation of the two cases:

(54)
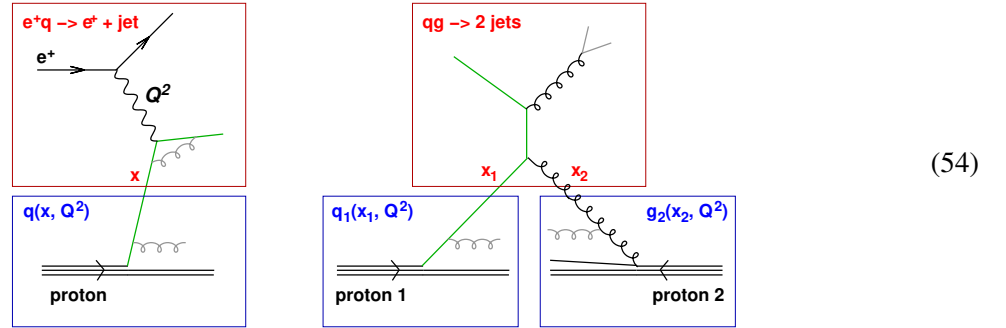
Given factorization and a determination of PDFs in DIS, one can simply take the expression Eq. (32) for a generic cross section in hadron–hadron collisions, and rewrite it with explicit factorization scales:

$$\sigma\sigma_{pp\to ZH} = \int dx_1 f_{q/p}(x_1, \mu_{\rm F}^2) \int dx_2 f_{\bar{q}/p}(x_2, \mu_{\rm F}^2)\, \hat{\sigma}_{q\bar{q}\to ZH}(x_1 p_1, x_2 p_2, \mu_{\rm F}^2).$$ (55)

Such a formula, with $\sigma_{q\bar{q}\to ZH}$ replaced by $\sigma_{q\bar{q}\to q\bar{q}}$ (summing also over processes with gluons, etc.) can be used to obtain predictions for the differential inclusive jet spectrum at the Tevatron. Figure 15 shows comparisons of data from DØ and from CDF with predictions from CTEQ6.5 [32] and MRST2004 [33] PDF parametrizations, illustrating excellent agreement. The two right-hand plots show how different incoming partonic scattering channels contribute to the cross section, highlighting the significant contribution from gluons.

It is perhaps misleading to use the word 'prediction' about Fig. 15: most advanced fully global PDF fits actually make use of data such as that in Fig. 15 as part of the fit. Still, it is a powerful consistency check that it is possible to obtain agreement both with the jet data, which is sensitive directly to quark and to gluon distributions, as well as with DIS data, which is directly sensitive to the quarks and indirectly to the gluons, through the scaling violations.

One technical comment is due concerning factorization. While our discussion has been limited to leading order, many of the figures that we have shown also involve higher orders (to which we will return). When using PDFs with predictions beyond LO, it is necessary to specify the 'factorization scheme', i.e., the specific procedure by which one separates emissions below and above $\mu_{\rm F}$. The figures in Sections 3.1 and 3.2 made use of the 'DIS' scheme (hence CTEQ6D), defined such that F_2 is given by Eq. (35), free of any $\mathcal{O}(\alpha_{\rm s})$ (or higher) corrections. While that scheme has the benefit of pedagogical simplicity, in real calculations (and all plots in this section) it is usually more convenient to use the '$\overline{\rm MS}$' factorization scheme, based on dimensional regularization.
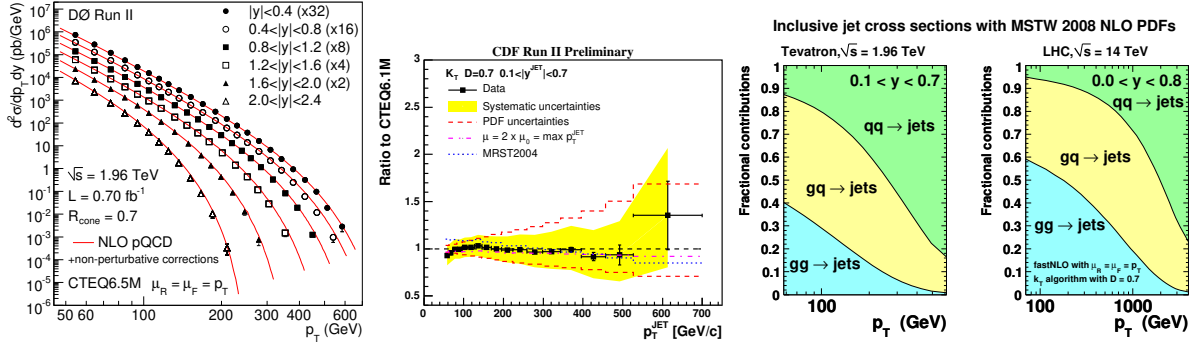
Fig. 15: Left: DØ inclusive jets data [34] compared to predictions with CTEQ6.5M PDFs. Middle: the ratio of CDF data from Ref. [35] to predictions with MRST2004 PDFs. Right: the relative contributions of different scattering channels to the Tevatron and LHC jet spectra, as a function of jet p_t, taken from Ref. [36].
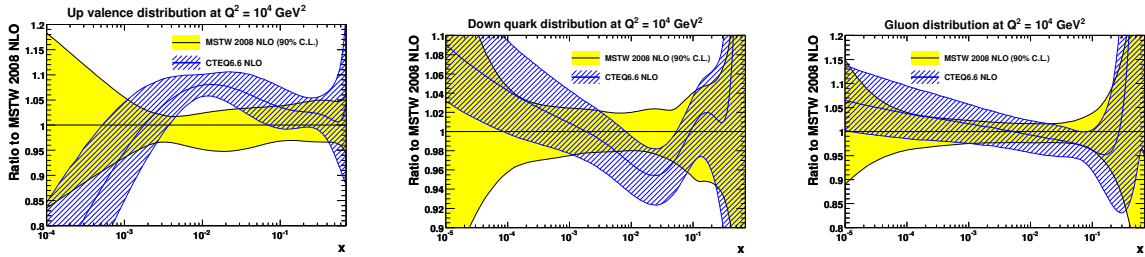
Fig. 16: Uncertainties on recent PDFs from the MSTW [37] and CTEQ groups [38] at a scale of $Q = 100\,\mathrm{GeV}$ (figure taken from Ref. [37]).
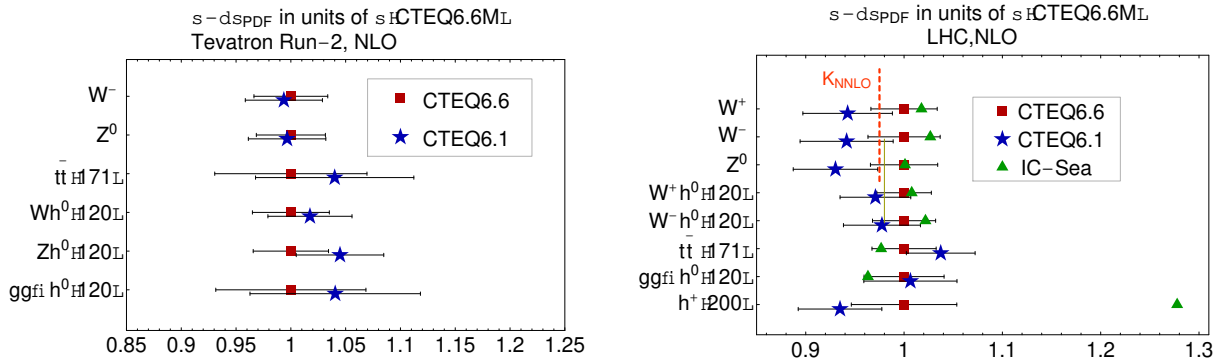
Fig. 17: Impact of PDF uncertainties on predictions for standard cross sections at the Tevatron and LHC from Ref. [38].

3.3.2 Uncertainties

An important part of the activity on global fits in recent years has been geared to the estimation of the uncertainties on PDFs. Figure 16 shows the uncertainties on recent PDF sets from the two most established PDF fitting groups, MSTW and CTEQ, illustrating uncertainties that are in the couple of per cent to ten per cent range. Figure 17 illustrates the impact of the uncertainties on predictions for a range of cross sections at the Tevatron and the LHC.

Estimating PDF uncertainties is something of an art: for example, one must parametrize the PDFs at some input scale and there is freedom in how flexible a parametrization one uses: too rigid (or with too many theorist's assumptions) and the global fit may not have the flexibility to describe the data or may

appear to have inappropriately small uncertainties in regions where there are no data; with too flexible a parametrization the fits may develop artefacts that adapt to statistical fluctuations of the data. Other issues include reconciling barely compatible data sets and deciding what values of χ^2 variations are reasonable to estimate errors given the incompatible data sets. In addition to MSTW and CTEQ (and several other groups), a recent entrant to PDF fitting is the NNPDF Collaboration, which is developing procedures that attempt to minimize theoretical bias in questions such as the parametrizations. First global fit results including pp data have been given in Ref. [39], though for fully accurate treatment of the HERA data, one should await their inclusion of heavy-quark effects. Their results so far tend to be similar to those of MSTW and CTEQ, except in the regions of small-x near the edge of the available HERA phasespace and for strange quark distributions, where they find somewhat larger uncertainties.

3.3.3 PDFs for LHC and the accuracy of DGLAP evolution

Figure 18 (left) illustrates the kinematic region in the x and Q^2 plane that is covered by the LHC (with $\sqrt{s} = 14\,\mathrm{TeV}$), compared to that for HERA and fixed-target experiments. The LHC region is labelled with a grid corresponding to mass (M, and one takes $Q = M$) and rapidity (y) of the object that is being produced. These are related to the incoming momentum fractions x_1 and x_2 and the pp squared centre-of-mass energy through

$$ M = \sqrt{x_1 x_2 s}\,, \qquad\qquad y = \frac{1}{2}\ln\frac{x_1}{x_2}\,. \qquad (56) $$

An object produced at a rapidity y involves $x_1 = \frac{M}{\sqrt{s}}e^{+y}$ and $x_2 = \frac{M}{\sqrt{s}}e^{-y}$. One feature that's immediately visible from the plot is that much of the LHC kinematic plane covers regions where PDFs have not been measured directly. We will therefore rely heavily on DGLAP evolution for our predictions. The right-hand plot in Fig. 18 illustrates just how much, in that it gives the factor by which the gluon distribution evolves in going from $Q = 2\,\mathrm{GeV}$ to $Q = 100\,\mathrm{GeV}$. Depending on the region in x, this can be a large factor, $\mathcal{O}(10)$. When compared to the experimental uncertainties on PDFs as shown in Figs. 16 and 17, we clearly have to ask how well we know the evolution.

In Section 4.1 we will discuss in detail how uncertainties are estimated in theoretical predictions. For now, essentially, there is freedom in Eq. (52) to choose a renormalization scale μ_R^2 for α_s that differs from μ_F^2. A conventional way of estimating the uncertainties is to choose $\mu_R^2 = (x_\mu \mu_F)^2$, varying x_μ in the range $\frac{1}{2} < x_\mu < 2$ (actually one often just takes three values, $x_\mu = \frac{1}{2}, 1, 2$). Starting from a fixed input at $2\,\mathrm{GeV}$ and evolving with different choices for x_μ gives the width of the band shown in Fig. 18 (right). That width is much larger than the uncertainties that we see in Figs. 16 and 17.

Fortunately we are not limited to leading-order (LO) DGLAP evolution, i.e., just the $\mathcal{O}(\alpha_s)$ term in Eq. (52). The order α_s^2 (next-to-leading order — NLO) corrections to the DGLAP equation were calculated around 1980 [41,42] and in 2004 the calculation of the NNLO corrections was completed [43, 44]. To give an idea of the complexity of that calculation, rather than taking a couple of lines as in Eq. (53) at LO, the NNLO results take $\mathcal{O}(10)$ pages to write down!

The impact of including higher orders on the evolution uncertainties is illustrated in Fig. 19. The same CTEQ61 input distribution is evolved from $2\,\mathrm{GeV}$ to $100\,\mathrm{GeV}$ with LO, NLO and NNLO splitting kernels, using three scale choices, $x_\mu = \frac{1}{2}, 1, 2$. The figure then shows the results for the gluon distribution at scale $100\,\mathrm{GeV}$, normalized to gluon distribution obtained with NNLO evolution and $x_\mu = 1$. One sees how the uncertainty comes down substantially as higher orders are included: 30% in some regions at LO, 5% over most of the x range at NLO and 2% at NNLO. In the NNLO case, the uncertainty is usually smaller than the experimental uncertainties for the PDFs that were shown in Fig. 16.[5]

---

[5]Figure 19 is to be interpreted with some caution. We have taken a fit carried out with $x_\mu = 1$, and then evolved it to high scales with $x_\mu \neq 1$. However, for example, the gluon distribution is partially determined from the evolution of F_2, so if the fit itself were carried out with $x_\mu \neq 1$, the fit result would probably change, introducing additional dependence of the $Q = 100$ results on the choice of x_μ. This could conceivably cancel some of the dependence seen in Fig. 19.
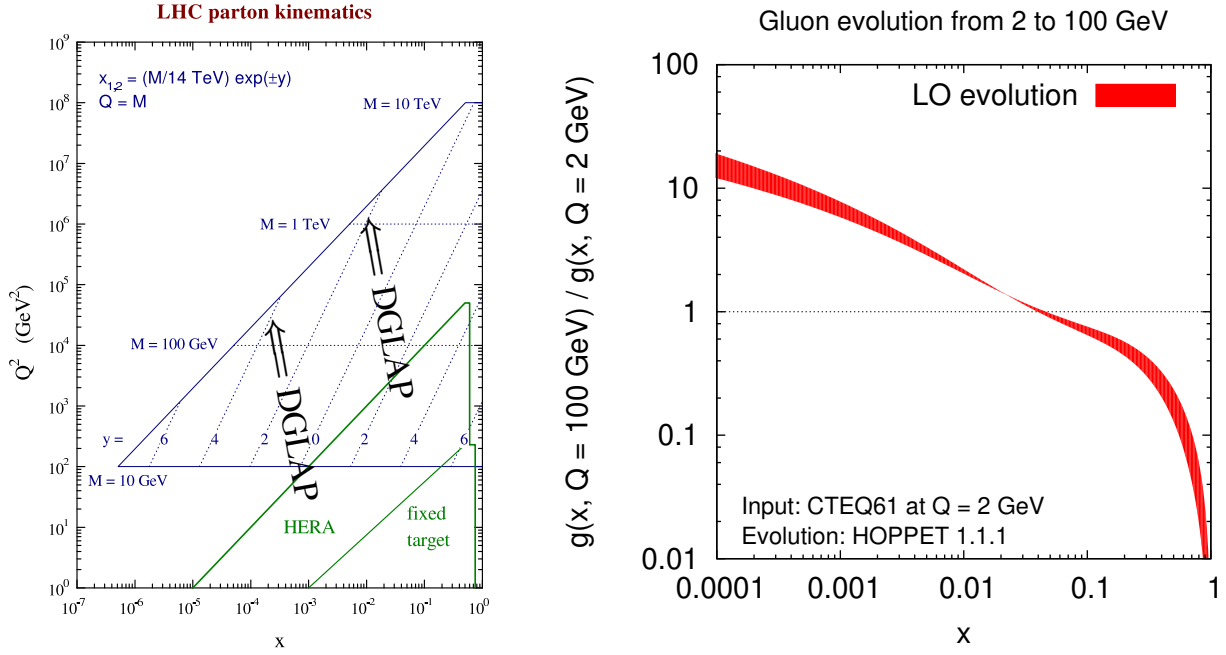
Fig. 18: Left: the kinematic regions covered by HERA, fixed-target scattering experiments, and by the LHC (adapted from the corresponding plot by Stirling). Right: the factor by which the gluon distribution evolves in going from a scale of $2\,\mathrm{GeV}$ to $100\,\mathrm{GeV}$ using CTEQ61 distributions as a fixed input at the low scale, and carrying out LO DGLAP evolution with HOPPET [40] with $x_\mu = \frac{1}{2}, 1, 2$ (see text for further details).
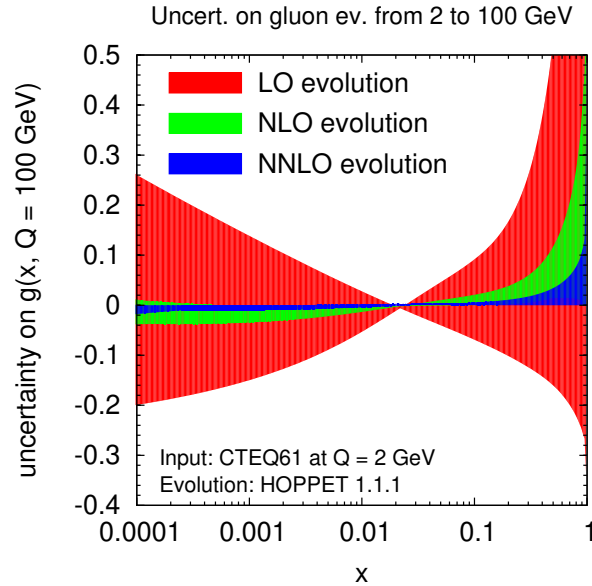
Fig. 19: Uncertainties on the evolution of the gluon distribution from a fixed (CTEQ61) input at scale $2\,\mathrm{GeV}$ up to $100\,\mathrm{GeV}$, for LO, NLO and NNLO evolution. The bands correspond to the envelope of the results with three scale choices $x_\mu = \frac{1}{2}, 1, 2$ and what is plotted is the ratio to the result at scale $100\,\mathrm{GeV}$, as obtained with NNLO evolution using $x_\mu = 1$.

Fig. 20: Left: production of a gluino ($\tilde{g}$) pair and the subsequent decay of each gluino through a squark ($\tilde{q}$) to quarks and a neutralino (χ^0); the exp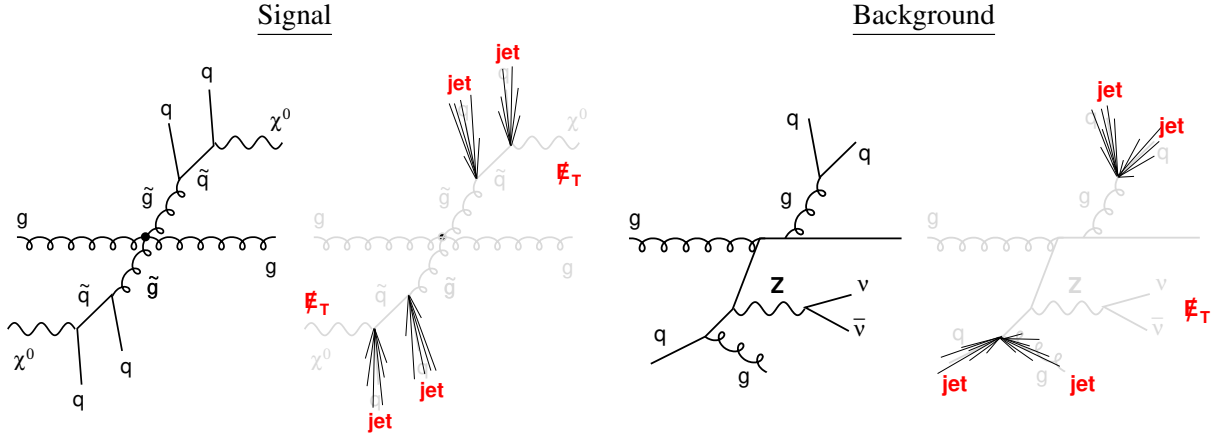erimental signature involves four jets and missing transverse energy ($\not{E}_T$) from the unobserved neutralino. Right: a background that mimics this signature, with missing energy coming from the production of a Z-boson that decays to neutrinos.

3.4 Summary

Here are some of the points to retain from this section. Firstly, the proton really is what we expect it to be, i.e., a uud state; however, fluctuations of the proton introduce many extra $q\bar{q}$ pairs ('sea'), as well as a substantial amount of 'glue', carrying 50% of the proton's momentum. The sea and gluon distributions diverge at small momentum fractions x.

Determination of the proton's PDFs involves fitting data from a range of experiments, with direct sensitivity to quarks (e.g., DIS), indirect sensitivity to the gluon (DIS Q^2 evolution), and direct sensitivity to quarks and gluons (jet data).

One of the keys to being able to measure consistent PDFs from different experiments, thinking about them in perturbative QCD and then applying them to predict results at new experiments is 'factorization': initial-state radiation, though collinear divergent, is process-independent; the divergent part can be absorbed into the definition of the PDFs, and then a universal set of PDFs, evolved between different scales with the DGLAP equations, can be used for any process.

Finally, the accuracy with which we know PDFs is quite remarkable: both from the experimental side and the theoretical side, in the best cases we know the PDFs to within a few per cent. This will be important in interpreting future signals of new particles, for example in Higgs-boson production at the LHC when we want to deduce its electroweak couplings given a measurement of its cross section.

If you need to use PDFs yourself, the best place to get them is from the LHAPDF library [45].

4 Predictive methods for LHC

In this section we will look at some of the different classes of technique can be used to make QCD predictions at LHC. Among the topics that we'll touch on are leading order (LO), next-to-leading order (NLO) and next-to-next-leading order (NNLO) calculations, parton-shower Monte Carlos, and then methods to combine the two types of calculation.

Many of the examples that we'll use will involve Z (and sometimes W) production at hadron colliders. One reason is that Z and W bosons decay to leptons and neutrinos (missing energy), both of which are easily-taggable handles that are characteristic of signals in many new-physics scenarios. An illustration is given in Fig. 20, which depicts supersymmetric production of a gluino pair and subsequent decay to four jets and missing transverse energy from the unobserved neutralinos. Because of the complexity of the decays and the fact that the missing energy is the sum of that from two neutralinos, it can be

27

difficult to extract clear kinematic structures (such as an invariant mass peak) that make a signal emerge unambiguously over the background. In such cases the contribution from the signal may just be to give a cross section that is larger than background expectations over a broad kinematic range. But that will only be a 'signal' if we understand what the backgrounds are.

The extent to which we will want to (or have to) rely on QCD predictions of backgrounds in deciding whether there are signals of new physics at the LHC is a subject that deserves in-depth consideration (for a nice discussion of it, see Ref. [46]). But QCD predictions will come into play in many other ways too. Monte Carlo parton shower programs, which simulate the full hadronic final state, are crucial in evaluating detector acceptances and response. And knowing QCD predictions (both for backgrounds and possible signals) is crucial in the design of methods to search for new physics, as well as for extracting meaning from the data (about couplings, spins, etc.) when, it is to be hoped, we finally see signals of something new.

4.1 Fixed-order predictions

Fixed-order predictions, which involve the first couple of terms in the QCD perturbative expansion for a given cross section, are conceptually quite simple: it is easy to state which contributions are included, and as one includes further orders in the expansion one can reasonably hope to see systematic improvement in the accuracy of one's predictions.

We'll first look at a couple of examples of fixed-order predictions, in order to develop a feel for how the perturbative expansion behaves, and how one estimates its accuracy. We will then examine more generally what theoretical inputs are needed for predictions for a given process, and what practical forms the predictive tools take.

4.1.1 Example 1: the cross section for $e^+e^- \to$ hadrons and its scale dependence

In Eq. (22), we wrote the total cross section for $e^+e^- \to$ hadrons as a perturbative series expansion in $\alpha_{\rm s}$ that multiplied the Born cross section $e^+e^- \to q\bar{q}$. The expansion was formulated in terms of the coupling evaluated at a renormalization scale $\mu_{\rm R}$ equal to the centre-of-mass energy Q, i.e., $\alpha_{\rm s}(\mu_{\rm R} = Q)$. That choice is, however, arbitrary: for example, the most energetic gluon that could be produced in $e^+e^- \to q\bar{q}g$ would be one with $E = Q/2$, so maybe we should be choosing $\mu_{\rm R} = Q/2$. And in loop diagrams, one integrates over gluon energies that go beyond Q, so maybe $\mu_{\rm R} = 2Q$ would be just as reasonable.

Because of this arbitrariness, a convention has emerged whereby one calculates a 'central value' for the prediction by setting the renormalization scale equal to the main physical scale for the process (e.g., the centre-of-mass energy at an e^+e^- collider; for hadron-collider processes the choice may be less obvious). The uncertainty is then estimated by varying the scale by a factor of two in either direction from the central value, i.e., taking $\frac{Q}{2} < \mu_{\rm R} < 2Q$. This is illustrated in Fig. 21 (left), which plots

$$\sigma^{\rm NLO} = \sigma_{q\bar{q}}(1 + c_1\alpha_{\rm s}(\mu_{\rm R})), \tag{57}$$

as a function of $\mu_{\rm R}$, showing how the $\mu_{\rm R}$-dependence translates into an uncertainty; note that c_1 can be read from Eq. (22). Given an expansion of the running coupling (i.e., of the middle result of Eq. (11), $\alpha_{\rm s}(\mu_{\rm R}) = \alpha_{\rm s}(Q) - 2b_0\alpha_{\rm s}^2(Q)\ln\frac{\mu_{\rm R}}{Q} + \mathcal{O}\left(\alpha_{\rm s}^3\right)$), we can rewrite Eq. (57) as

$$\sigma^{\rm NLO}(\mu_{\rm R}) = \sigma_{q\bar{q}}\left(1 + c_1\alpha_{\rm s}(Q) - 2c_1b_0\alpha_{\rm s}^2(Q)\ln\frac{\mu_{\rm R}}{Q} + \mathcal{O}\left(\alpha_{\rm s}^3\right)\right). \tag{58}$$

This tells us that as we vary the renormalization scale for a prediction up to $\mathcal{O}\left(\alpha_{\rm s}\right)$ (NLO), we effectively introduce $\mathcal{O}\left(\alpha_{\rm s}^2\right)$ (NNLO) pieces into the calculation: by generating some fake set of NNLO terms, we are probing the uncertainty of the cross section associated with the missing full NNLO correction.
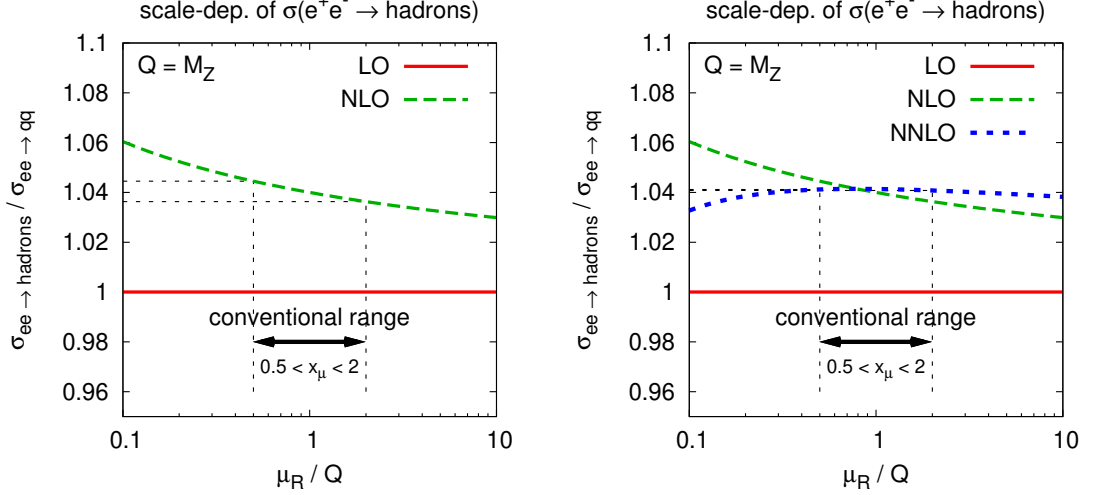
Fig. 21: Renormalization-scale dependence of the NLO (left) and NNLO (right) predictions for the $e^+e^- \to$ hadrons total cross section, together with an indication of the conventional choice of scale-variation range

If we calculate the actual NNLO cross section for general μ_R, it will have a form

$$\sigma^{\text{NNLO}}(\mu_R) = \sigma_{q\bar{q}} \left(1 + c_1 \alpha_s(\mu_R) + c_2(\mu_R)\alpha_s^2(\mu_R)\right) . \tag{59}$$

Observe that the c_2 coefficient now depends on μ_R. This is necessary because the second-order coefficient must cancel the $\mathcal{O}\left(\alpha_s^2\right)$ ambiguity due to the scale choice in Eq. (58). This constrains how $c_2(\mu_R)$ depends on μ_R:

$$c_2(\mu_R) = c_2(Q) + 2c_1 b_0 \alpha_s^2(Q) \ln \frac{\mu_R}{Q} , \tag{60}$$

where $c_2(Q)$ can again be read from Eq. (22). If we now express $\sigma^{\text{NNLO}}(\mu_R)$ in terms of $\alpha_s(Q)$, we will find that the residual dependence on μ_R appears entirely at $\mathcal{O}\left(\alpha_s^3\right)$, i.e., one order further than in Eq. (58). This is reflected in the right-hand plot of Fig. 21, which illustrates how the impact of the scale variation at NNLO is significantly reduced, since we are now probing the impact of missing α_s^3 terms, rather than α_s^2 terms.

If we had an arbitrarily large number of terms in the α_s expansion, the scale dependence would disappear exactly. The fact it doesn't in the presence of a fixed number of terms may initially seem like a drawback, but in some respects it's a blessing in disguise because it provides a useful handle on the uncertainties. This is why scale variation has become a standard procedure. It's worth bearing in mind that it isn't a failsafe mechanism: a trivial example comes from the LO curve in Fig. 21. It doesn't have any scale variation because they don't depend on α_s, yet it differs significantly from the higher-order results.

4.1.2 Example 2: $pp \to Z$

At LO the $pp \to Z$ cross section involves a single underlying hard partonic process, namely $q\bar{q} \to Z$, which is purely electroweak. To go from the $q\bar{q} \to Z$ squared matrix element to the $pp \to Z$ result, one must integrate over the quark distributions

$$\sigma^{\text{LO}}_{pp \to Z} = \sum_i \int dx_1 dx_2 \, f_{q_i}(x_1, \mu_F^2) \, f_{\bar{q}_i}(x_2, \mu_F^2) \, \hat{\sigma}_{0, q_i \bar{q}_i \to Z}(x_1 p_1, x_2 p_2) , \tag{61}$$

for which one must choose a factorization scale μ_F. A natural choice for this scale is $\mu_F = M_Z$, but as with the renormalization scale it is conventional to vary it by a factor of two either side of the central choice in order to obtain a measure of the uncertainties in the prediction.

29

Adding NLO and NNLO terms, the structure becomes

$$\sigma_{pp \to Z+X}^{\text{NNLO}} = \sum_{i,j} \int dx_1 dx_2 \, f_i(x_1, \mu_{\text{F}}^2) \, f_j(x_2, \mu_{\text{F}}^2) \bigg[\hat{\sigma}_{0,ij \to Z}(x_1, x_2) \, + \, \alpha_{\text{s}}(\mu_{\text{R}}) \hat{\sigma}_{1,ij \to Z+X}(x_1, x_2, \mu_{\text{F}})$$
$$+ \, \alpha_{\text{s}}^2(\mu_{\text{R}}) \hat{\sigma}_{2,ij \to Z+X}(x_1, x_2, \mu_{\text{F}}, \mu_{\text{R}}) \bigg]. \quad (62)$$

We now have a sum over the flavours i and j of the initial partons, because starting from NLO there are contributions from (say) gluon-quark scattering [cf. Fig. 22 (left)]. The cross section is written as being for $Z+X$, where the X means that we allow anything (e.g., quarks, gluons) to be produced in addition to the Z-boson. At $\mathcal{O}(\alpha_{\text{s}})$ the μ_{F} dependence of the σ_1 coefficient partially cancels the dependence present at $\mathcal{O}(\alpha_{\text{s}}^0)$ coming from the μ_{F} dependence of the PDFs. That dependence is further cancelled at $\mathcal{O}(\alpha_{\text{s}}^2)$, as is part of the μ_{R} dependence that is introduced in the $\mathcal{O}(\alpha_{\text{s}}(\mu_{\text{R}}))$ term. The plot on the right of Fig. 22 shows the Z-boson cross section as a function of its rapidity [47]. The bands indicate the uncertainty due to scale variation (taking $\frac{1}{2} M_Z < \mu_{\text{R}} = \mu_{\text{F}} < 2 M_Z$)[6] and show how this uncertainty undergoes important reductions going from LO to NLO to NNLO.

One of the interesting features that comes out of Fig. 22 is that the LO prediction is only good to within a factor of 1.5 to 2, despite the fact that $\alpha_{\text{s}}(M_Z) \simeq 0.118$ would imply 10% accuracy. This is because the $\mathcal{O}(\alpha_{\text{s}})$ corrections come with large coefficients. This is not uncommon in hadron-collider cross sections. Furthermore the LO uncertainty band seems not to provide a faithful measure of the true uncertainty. Other aspects of the perturbative expansion do seem to behave as one would expect: the size of the uncertainty band decreases significantly going from LO to NLO (10–20%) to NNLO (a few per cent). And the actual shift in the central value in going from NLO to NNLO is substantially smaller than that from NLO to LO.

Are these characteristics representative of the 'typical' situation for collider observables? We only have predictions up to NNLO in a handful of cases (see below) and in those it is. In cases where we just have NLO predictions, the features of large 'K-factors' (NLO/LO enhancements) with a reduced NLO uncertainty band are not uncommon, suggesting that beyond NLO corrections should be small. Exceptions are known to arise in two types of case: those where new enhanced partonic scattering channels open up at NLO (or beyond); and that involve two disparate physical scales. For example, if you ask for the Z-boson to have a transverse momentum p_t that is much smaller than M_Z, then each power of α_{s} in the expansion of the cross section will be accompanied by up to two powers of $\ln^2 M_Z/p_t$, leading to large coefficients at all orders in the perturbative expansion. These are due to incomplete cancellation between real and virtual (loop) divergences: loop corrections do not affect the Z-boson p_t and so are fully integrated over, whereas real emissions do affect the Z p_t and so are only partially integrated over.

4.1.3 Predictions for more complex processes

As an example of a more complex process, consider the production of a Z-boson plus a jet. The leading order cross section requires the calculation of the $\mathcal{O}(\alpha_{\text{s}})$ squared diagrams for $q\bar{q} \to Zg$, $qg \to Zq$ and $\bar{q}g \to Z\bar{q}$. The NLO cross section additionally requires all $\mathcal{O}(\alpha_{\text{s}}^2)$ contributions with a Z boson and at least one jet, as illustrated in Fig. 23, i.e., the squared tree-level diagram for $ij \to Z + 2$ partons and the interference of the 1-loop and tree-level diagrams for $ij \to Z + 1$ parton.

More generally, Fig. 23 allows you to read off the contributions that you will need for an N^pLO calculation of $ij \to Z + $n partons: just take all entries in Fig. 23 with at least n partons, up to order α_{s}^{n+p}. Entries in black are known and have already been used to obtain predictions for LHC and Tevatron. The one entry in grey, the 2-loop $Z + 1$ parton contribution, is known but has yet to be used in any hadron-collider prediction, for reasons that we will discuss below. Entries that are absent (e.g., $Z + 2$ partons at two loops) have so far proven to be too complicated to calculate.

---

[6]The variation of μ_{R} and μ_{F} simultaneously, though common, is not the only possible procedure. An attractive alternative is to vary both independently around a central scale, with the additional requirement that $\frac{1}{2} < \mu_{\text{R}}/\mu_{\text{F}} < 2$ [48].
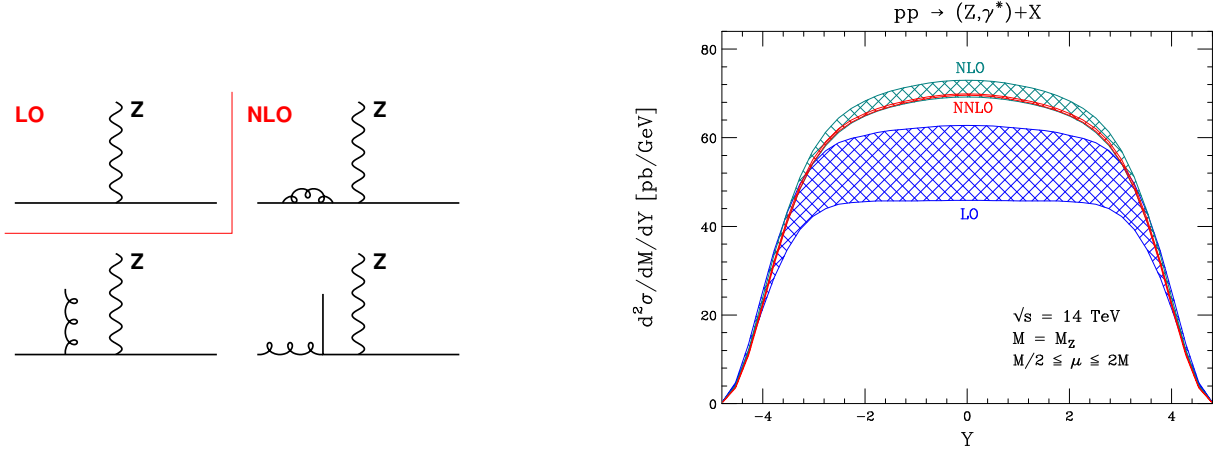
Fig. 22: Left: classes of diagram that appear for $pp \to Z$ at LO and at NLO. Right: cross section at the LHC for the Z-boson, differential in rapidity, at LO, NLO and NNLO, as taken from Ref. [47].
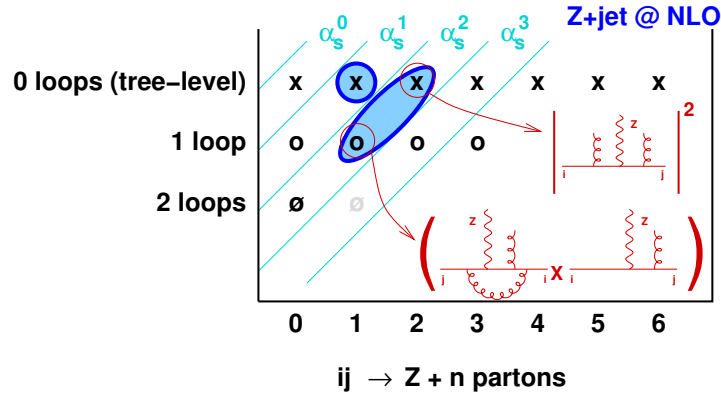

Fig. 23: Illustration of the contributions that are known for $ij \to Z + $ n partons, where i and j are arbitrary incoming partons, according to the number of outgoing partons, the number of loops and the number of powers of the coupling. An 'x' represents a squared tree-level diagram, an 'o' represents the interference of a 1-loop diagram with a tree-level diagram, and a 'ø' represents the interference of a two-loop diagram with a tree-level diagram or the square of a 1-loop diagram. Entries in black are known and used; entries in grey are known but have not been used. The entries in the shaded ellipses are those that are relevant for the NLO calculation of the cross section for the production of a Z-boson with a jet.

The classes of contributions calculated for $ij \to Z + $ n partons provide a representative view of the situation for other processes as well, with tree-level diagrams calculated up to quite high final-state multiplicities, ~ 10, 1-loop diagrams having been used for processes with up to 3 or sometimes 4 final-state particles, and 2-loop diagrams available and used only for $2 \to 1$ type processes, essentially $pp \to W$, $pp \to Z/\gamma^*$ and $pp \to H$.

It's natural to ask in what form these various calculations are available. For certain very simple quantities, for example, the total cross section for $t\bar{t}$ production, or for W, Z or Higgs production, the result of the perturbative calculation can be written as in Eq. (62),

$$\sigma^{\text{N}^p\text{LO}}_{pp \to A+X} = \sum_{i,j} \int dx_1 dx_2\, f_i(x_1, \mu_F^2)\, f_j(x_2, \mu_F^2) \times \sum_{m=0}^{p} \alpha_s^{n+m}(\mu_R)\, \hat{\sigma}_{m,ij \to A+X}(x_1 x_2 s, \mu_R, \mu_F),$$

(63)

where the $\sigma_{m,ij \to A+X}(x_1 x_2 s, \mu_R, \mu_F)$ are functions whose analytical expressions can be found in the relevant papers (Refs. [49, 50] for W and Z, and Refs. [50–53] for Higgs-boson production). To obtain

31

Table 1: The number of Feynman diagrams for tree level $gg \to N$ gluon scattering [59]

N	2	3	4	5	6	7	8
No. diags	4	25	220	2485	34300	5×10^5	10^7

a prediction, one just has to type them into a computer program and then integrate over x_1 and x_2.

In most cases, however, one wants to calculate a cross section that incorporates experimental cuts, such as lepton acceptances, transverse momentum cuts on jets, etc. In these cases the type of tool that should be used depends on the order to which you want the answer.

4.1.3.1   LO predictions

As long as one is dealing with infrared safe observables, then for a LO prediction one need only include tree-level diagrams, in kinematic regions in which their contributions are finite. The simplest approach therefore is to carry out Monte Carlo integration over phase-space points, have a subroutine that determines whether a given phase-space point passes the cuts, and if it does calculate the squared matrix elements and PDF factors for each possible partonic subprocesses.

Quite a number of tools enable you to do this: ALPGEN [54], COMIX/SHERPA [55], COM-PHEP [56], HELAC/PHEGAS [57] and MADGRAPH [58]. They allow you to calculate cross sections for a broad range of $2 \to n$ scattering processes with n up to 6–8 (or in some cases even beyond). Some of these (COMPHEP, MADGRAPH) use formulae obtained from direct evaluations of Feynman diagrams. This gives them significant flexibility in terms of the range or processes they support (e.g., with easy inclusion of many new physics models), though they suffer at large n because the number of diagrams to evaluate grows rapidly as n increases (cf. Table 1). Others (ALPGEN, HELAC/PHEGAS and COMIX/SHERPA) use methods designed to be particularly efficient at high multiplicities, such as Berends–Giele recursion [60], which builds up amplitudes for complex processes by recursively reusing simpler ones (a nice technical review of the technique is given in Ref. [61]).

4.1.3.2   NLO predictions

When, in Sections 2.2 and 2.3.2, we looked at the cancellation between divergences in real and loop diagrams, we wrote the loop diagram with an explicit integral over phase space so as to be able to match the divergences between real and loop diagrams and cancel them easily.

A subtlety that we ignored is that in practical evaluations of loop diagrams, the integral over loop momenta is carried out in $4 - \epsilon$ dimensions rather than 4 dimensions, in order to regularize the divergences that appear and obtain an answer whose finite part is known independently of any related tree-level diagrams. On the other-hand, experimental cuts are defined in four dimensions, so the real tree-level diagrams must be integrated in four dimensions, which implies divergent results if the real diagrams are taken alone.

This mismatch between the ways loop and tree-level diagrams are handled is one of the main difficulties in carrying out calculations that include experimental cuts beyond LO. For the calculation of a process with n partons at LO, the standard technique nowadays to deal with the problem is to introduce a $n+1$-parton *counterterm* where the $n+1^{\text{th}}$ parton is always soft and collinear so that it doesn't affect an IR safe observable. It is subtracted from the $n+1$-parton real diagram in four dimensions and designed so as to cancel all of its soft and collinear divergences. It is also designed such that the kinematics of its $n+1^{\text{th}}$ parton can be integrated analytically in $4 - \epsilon$ dimensions so that the result can be easily added to the loop diagram and cancel *its* divergences. Since the counterterm is subtracted once and added once, its net impact is null, and can just be thought of as a way of reshuffling divergences. This is known as a 'subtraction' procedure and the variant most widely used in current NLO computer codes is 'dipole'
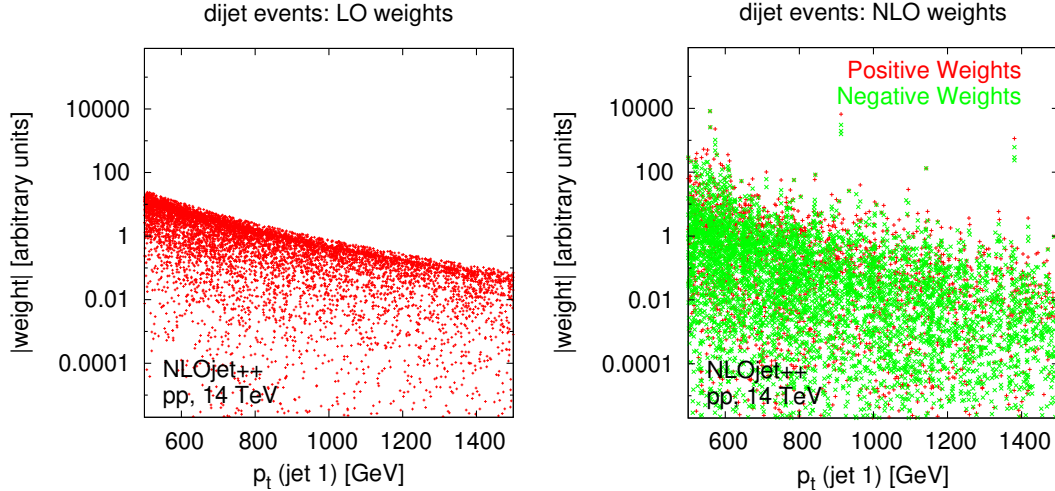
Fig. 24: Left: weights associated with tree-level LO $2 \to 2$ scattering events in the calculation of the dijet cross section, shown as a function of the transverse momentum of the harder jet. Right: weights at NLO from real $2 \to 3$ events, subtraction counterterm events and loop events (after addition of the integrated counterterm), again as a function of p_t of the hardest jet. Results obtained with NLOjet++ [65].

subtraction [62]; other methods that have seen numerous applications include FKS [63] and antenna [64] subtraction.

An illustration of how subtraction works in practice is given in Fig. 24 for dijet production. In the left-hand plot we see weights for the LO process as a function of the jet p_t. Though there is some dispersion in the weights, there is a clear upper bound as a function of p_t, reflecting the finiteness of the matrix-elements. The right-hand plot shows weights at NLO. Here there is no clear upper bound to the weights; however we do see that unusually large (nearly divergent) weights come in clusters: usually one positive (red, '+') weight, accompanied by one or more negative (green, '×') weights with identical jet transverse momenta, so that each event in the cluster contributes to the same bin of the cross section and their weights sum to a finite result.

Technically, one main consideration has so far limited the range of processes for which NLO results exist: the availability of the loop amplitude. Until recently loop amplitudes were usually calculated semi-manually for each process. The complexity of the calculations increased significantly with the number of outgoing legs, limiting available results to those with at most three outgoing partons. Many NLO results for $2 \to 2$ and $2 \to 3$ processes are incorporated into programs such as NLOJET++ for jet production [65], MCFM for processes with heavy quarks and/or heavy electroweak bosons [66], VBFNLO for vector-boson fusion processes [67], and the PHOX family [68] for processes with photons in the final state.

In the past couple of years, techniques have come to fruition that offer the prospect of automated calculation of arbitrary loop diagrams. Though full automation is not here yet, a number of $2 \to 4$ processes have now been calculated at NLO thanks to these advances, including $pp \to t\bar{t}b\bar{b}$ [69, 70], $pp \to t\bar{t}jj$ [71] (where j represents a jet) and $pp \to W+3j$ [72, 73] and $pp \to Z+3j$ [74].[7]

It should be said that NLO calculations are very computing intensive: for some observables it is not unusual to have to devote several years of CPU time in order to get adequate numerical convergence of the Monte Carlo integration.

---

[7]Since the original version of this writeup, first results for a $2 \to 5$ process have also appeared [75].

4.1.3.3 NNLO predictions

NNLO predictions suffer from the same problem of cancelling divergences between real and virtual corrections that is present at NLO, with the complication that instead of having one soft and one collinear divergence, there are now two of each, greatly complicating the task of figuring out counterterms to allow experimental cuts to be implemented in four dimensions.

As a result, the only general subtraction type approaches that exist currently are for processes without incoming hadrons, notably $e^+e^- \to 3j$ [76,77]. For hadron collider processes it is only $2 \to 1$ processes that are available, specifically vector-boson (FEWZ [78] and DYNNLO [79]) and Higgs-boson (FEHIP [80] and HNNLO [81]) production, using methods that are not so easily generalizable to more complicated processes.

4.2 Monte Carlo parton-shower programs

The programs we've discussed so far, known as 'Matrix Element Monte Carlos' provide a powerful combination of accuracy and flexibility as long as you want to calculate IR and collinear safe observables (jets, W's, Z's, but not pions, kaons, etc.), don't mind dealing with wildly fluctuating positive and negative weights, and don't need to study regions of phase space that involve disparate physical scales.

All these defects are essentially related to the presence of soft and collinear divergences. Yet we know that real life does not diverge. So it is natural to wonder whether we can reinterpret the divergences of perturbation theory physically. It turns out that the right kind of question to ask is "what is the probability of *not* radiating a gluon above some (transverse momentum) scale k_t". Starting from a $q\bar{q}$ system, using the results of Section 2, we know that to $\mathcal{O}(\alpha_s)$ the answer in the soft and collinear limit goes as

$$P(\text{no emission above } k_t) \simeq 1 - \frac{2\alpha_s C_F}{\pi} \int^Q \frac{dE}{E} \int^{\pi/2} \frac{d\theta}{\theta} \Theta(E\theta - k_t). \tag{64}$$

It so happens that in the soft and collinear limit, this result is easy to work out not just at first order, but at all orders, giving simply the exponential of the first order result

$$P(\text{no emission above } k_t) \equiv \Delta(k_t, Q) \simeq \exp\left[-\frac{2\alpha_s C_F}{\pi} \int^Q \frac{dE}{E} \int^{\pi/2} \frac{d\theta}{\theta} \Theta(E\theta - k_t)\right]. \tag{65}$$

Whereas Eq. (64) had a 'bare' infinity if one took $k_t \to 0$, Eq. (65) is simply bounded to be between 0 and 1.

The quantity $\Delta(k_t, Q)$ is known as a Sudakov form factor. We've been very approximate in the way we've calculated it, neglecting for example the running of the coupling (α_s should be placed inside the integral and evaluated at the scale $E\theta$) and the treatment of hard collinear radiation (the dE/E integral should be replaced with the full collinear splitting function), but these are just technical details. The importance of the Sudakov form factor is that it allows us to easily calculate the distribution in transverse momentum k_{t1} of the gluon with largest transverse momentum in an event:

$$\frac{dP}{dk_{t1}} = \frac{d}{dk_{t1}} \Delta(k_{t1}, Q). \tag{66}$$

This distribution is easy to generate by Monte Carlo methods: take a random number r from a distribution that's uniform in the range $0 < r < 1$ and find the k_{t1} that solves $\Delta(k_{t1}, Q) = r$. Given k_{t1} we also need to generate the energy for the gluon, but that's trivial. If we started from a $q\bar{q}$ system (with some randomly generated orientation), then this gives us a $q\bar{q}g$ system. As a next step one can work out the Sudakov form factor in the soft/collinear limit for there to be no emission from the $q\bar{q}g$ system as a whole above some scale k_{t2} ($< k_{t1}$) and use this to generate a second gluon. The procedure is then repeated over and over again until you find that the next gluon you would generate is below some non-perturbative cutoff scale Q_0, at which point you stop. This gives you one 'parton shower' event.

a)

```
                                     ---INITIAL STATE---

IHEP   ID    IDPDG IST MO1 MO2 DA1 DA2   P-X      P-Y      P-Z    ENERGY    MASS
  1 P         2212 101   0   0   0   0   0.00     0.00   7000.0   7000.0    0.94
  2 P         2212 102   0   0   0   0   0.00     0.00  -7000.0   7000.0    0.94
  3 CMF          0 103   1   2   0   0   0.00     0.00      0.0  14000.0   14000.0
```

b)

```
                                    ---HARD SUBPROCESS---

IHEP   ID    IDPDG IST MO1 MO2 DA1 DA2   P-X      P-Y      P-Z    ENERGY    MASS
  4 UQRK         2 121   6   8   9   5    0.00     0.00    590.8    590.8    0.32
  5 GLUON       21 122   6   4  17   8    0.00     0.00   -232.1    232.1    0.75
  6 HARD         0 120   4   5   7   8    0.40    -9.40    358.7    823.0   740.63
  7 ZO/GAMA*    23 123   6   7  22   7 -261.59  -217.31    329.3    481.6    88.56
  8 UQRK         2 124   6   5  23   4  261.59   217.31     29.4    341.3    0.32
```

c)

```
                                     ---PARTON SHOWERS---

IHEP   ID    IDPDG IST MO1 MO2 DA1 DA2   P-X      P-Y      P-Z    ENERGY    MASS
  9 UQRK        94 141   4   6  11  16    2.64    -9.83    592.2    590.2   -49.07      ┐
 10 CONE         0 100   4   5   0   0   -0.27     0.96      0.1      1.0     0.00      │
 11 GLUON       21   2   9  12  32  33   -1.02     3.59      5.6      6.7     0.75-     │ INITIAL
 12 GLUON       21   2   9  13  34  35    0.25     1.46      3.6      4.0     0.75-     │ STATE
 13 GLUON       21   2   9  14  36  37   -0.87     1.62      4.7      5.1     0.75-     │ SHOWER
 14 GLUON       21   2   9  15  38  39   -0.81     4.17   3611.7   3611.7     0.75-     │
 15 GLUON       21   2   9  16  40  41   -0.19    -1.01   1727.7   1727.7     0.75-     │
 16 UD        2101   2   9  25  42  41    0.00     0.00   1054.6   1054.6     0.32-     │
 17 GLUON       94 142   5   6  19  21   -2.23     0.44   -233.5    232.8   -18.36      │
 18 CONE         0 100   5   4   0   0    0.77     0.64      0.2      1.0     0.00      │
 19 GLUON       21   2  17  20  43  44    1.60     0.58     -2.1      2.8     0.75      ┘
 20 UD        2101   2  17  21  45  44    0.00     0.00  -2687.6   2687.6     0.32      ┐
 21 UQRK         2   2  17  32  46  45    0.63    -1.02  -4076.9   4076.9     0.32      │
 22 ZO/GAMA*    23 195   7  22 251 252 -257.66  -219.68    324.8    477.5    88.56      │
 23 UQRK        94 144   8   6  25  31  258.06   210.29     33.9    345.5    86.10      │ FINAL
 24 CONE         0 100   8   5   0   0    0.21     0.17     -1.0      1.0     0.00      │ STATE
 25 UQRK         2   2  23  26  47  42   26.82    24.33     23.7     43.3     0.32      │ SHOWER
 26 GLUON       21   2  23  27  48  49    8.50     8.18      6.0     13.3     0.75      │
 27 GLUON       21   2  23  28  50  51   73.27    61.24     12.0     96.2     0.75      │
 28 GLUON       21   2  23  29  52  53   73.66    58.54     -6.3     94.3     0.75      │
 29 GLUON       21   2  23  30  54  55   67.58    52.13     -7.3     85.7     0.75      │
 30 GLUON       21   2  23  31  56  57    6.98     4.60      2.3      8.7     0.75      │
 31 GLUON       21   2  23  43  58  59    1.24     1.26      3.6      4.1     0.75      ┘
```

Fig. 25: Sequence of steps in the generation of of a $pp \to Z+j$ event in HERWIG: (a) specification of the colliding beams and their energy, (b) generation of the kinematics and partonic flavour of the hard subprocess, $ug \to Zu$, and (c) generation of the initial- and final-state parton showers.

This is essentially the procedure that's present in the shower of PYTHIA 8 [82] and the p_t ordered option of PYTHIA 6.4 [83], as well as ARIADNE [84] and SHERPA 1.2 (the SHERPA reference [85] describes an earlier version). It is also possible to choose other ordering variables: the original PYTHIA shower [83, 86] is based on virtuality ordering (plus an angular veto). This is still the most widely used shower which works well on a range of data (though there are theoretical issues in some formulations of virtuality ordering). In the HERWIG family of programs [87, 88] it is angular ordering that is used.

The above shower descriptions hold for final-state branching. With initial-state hadrons, one also needs to be careful with the treatment of the PDFs, since the collinear splitting that is accounted for in the parton shower is also connected with the way the PDF is built up at the scale of the hard scattering.

The sequence of steps for the generation of a parton-shower event in pp collisions is illustrated in Fig. 25.

Real events consist not of partons but of hadrons. Since we have no idea how to calculate the transition between partons and hadrons, Monte Carlo event generators resort to 'hadronization' models. One widely-used model involves stretching a colour 'string' across quarks and gluons, and breaking it up into hadrons [89, 90]. For a discussion of the implementation of this 'Lund' model in the MC program PYTHIA, with further improvements and extensions, Ref. [86] and references therein provide many details. Another model breaks each gluon into a $q\bar{q}$ pair and then groups quarks and anti-quarks into colourless 'clusters', which then give the hadrons. This cluster type hadronization is implemented in the HERWIG event generator [87, 88, 91] and recent versions of SHERPA. Both approaches are illustrated in Fig. 26

Hadronization models involve a number of 'non-perturbative' parameters. The parton-shower itself involves the non-perturbative cutoff Q_0. These different parameters are usually tuned to data from the LEP experiments. The quality of the description of the data that results is illustrated in Fig. 27.

A final point in our brief description of Monte Carlo event generators concerns the 'underlying event' in pp and γp collisions. In addition to the hard interaction that is generated by the Monte Carlo simulation, it is also necessary to account for the interactions between the incoming proton (or photon) remnants. This is usually modelled through multiple extra $2 \to 2$ scattering, occurring at a scale of a few GeV, and known as multiple parton interactions. This modelling of the underlying event is crucial in order to give an accurate reproduction of the (quite noisy) energy flow that accompanies hard scatterings in hadron-collider events.

Our description here of Monte Carlo event generators has been fairly brief. For a more complete discussion, a good starting point is the lectures notes by Sjöstrand [93] from the 2006 School.
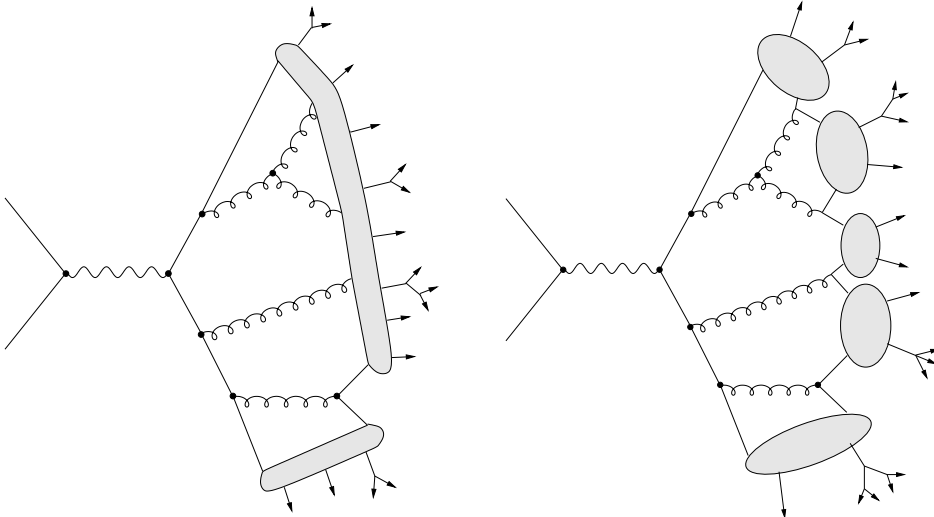
Fig. 26: Illustration of string (left) and cluster (right) fragmentation, taken from Ref. [1]
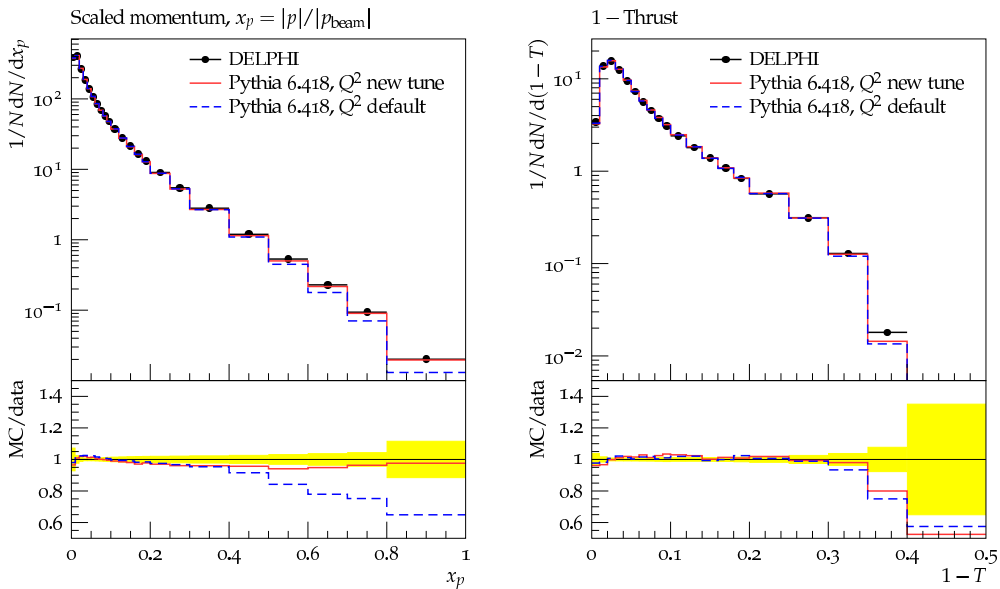

Fig. 27: A comparison of DELPHI e^+e^- data for the particle (scaled) momentum distribution (left) and the thrust event shape distribution (right) with two tunes of the PYTHIA event generator [92].

4.3 Comparing fixed-order and parton-shower programs

Parton-shower Monte Carlo programs do a good job of describing most of the features of common events, including the hadron-level detail that is essential for the correct simulation of detector effects on event reconstruction. Another nice feature of theirs is that events have equal weight, just as with real data.

A drawback of parton-shower Monte Carlos is that, because they rely on the soft and collinear approximation, they do not necessarily generate the correct pattern of hard large-angle radiation. This can be important, e.g., if you're simulating backgrounds to new-physics processes, for which often the rare, hard multi-jet configurations are of most interest. In contrast, fixed-order programs do predict these configurations correctly.

The purpose of this section is to give two examples of comparisons between parton-shower predictions and fixed-order predictions, in order to help illustrate their relative strengths.
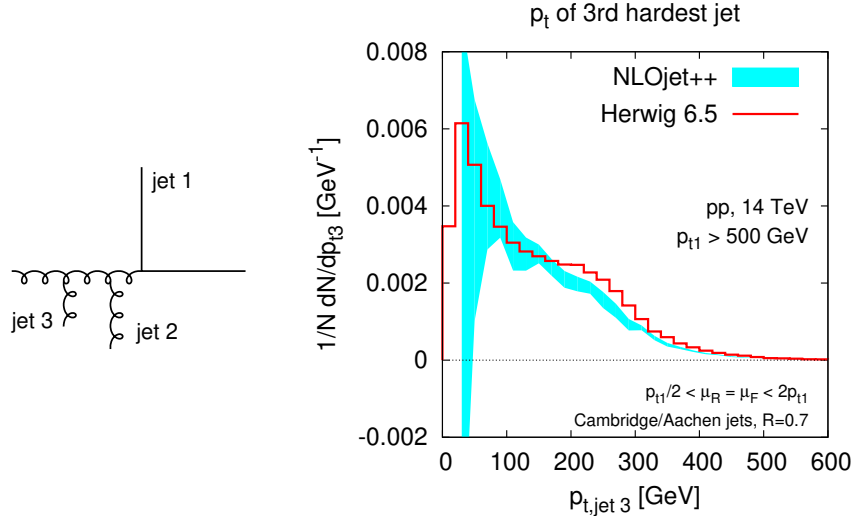
Fig. 28: The predicted p_t distribution for the jet with the third largest transverse momentum, p_{t3}, in $14\,\text{TeV}$ pp events where the hardest jet has transverse momentum $p_{t1} > 500\,\text{GeV}$. The solid histogram is the result from HERWIG 6.5 [87]. The band corresponds to the ratio of the NLO 3-jet cross section, differential in p_{t3} (with the p_{t1} cut), to the NLO cross section for events to pass the p_{t1} cut. The width of the band corresponds to the scale uncertainty and the results have been obtained with NLOJET++ [65].

4.3.1 Jet production

In plain jet production, parton shower Monte Carlos start from a hard event that consists of parton–parton scattering diagrams like $qq \to qq$, $qg \to qg$, etc., and then rely on the showering to generate extra radiation. While the showering is only correct in the soft and/or collinear limit, it does sometimes generate extra hard radiation. In Fig. 28 we can see distribution in transverse momentum of the 3rd hardest jet, in HERWIG events in which we have imposed a cut on the hardest jet, $p_{t1} > 500\,\text{GeV}$. That is compared to the NLO (NLOJET++) prediction for the same distribution, together with its uncertainty band from scale variation.

In much of the range, one observes relatively good agreement between the two distributions: the 20–30% differences that are visible around $p_{t3} \equiv p_{t,\text{jet}\,3} \sim 250\,\text{GeV}$ are no larger than the uncertainties that would be obtained from a LO calculation, despite the fact that in this region HERWIG does not even include the exact LO $2 \to 3$ matrix element. Of course, it is hard to be sure whether the good agreement is meaningful generally, or instead just a coincidence specific to our particular choice of observable — and the only way to be sure is, for each new observable, to also generate the NLO prediction.

The NLO prediction is not without its limitations though: at low p_{t3}, the uncertainty band on the NLO prediction blows up. This is a manifestation of large higher-order corrections, which compromise the convergence of the perturbative series. They arise because we have a large ratio of scales between $p_{t1}(\gtrsim 500\,\text{GeV})$ and p_{t3} (a few tens of GeV). Such large scale ratios translate into NLO corrections whose size relative to the LO contribution go as $\alpha_s \ln^2 p_{t1}/p_{t3} \sim 1$ and $\alpha_s \ln p_{t1}/p_{t3}$.

4.3.2 Vector-boson plus production

The picture seen above of good agreement between parton shower Monte Carlo and fixed-order predictions does not always hold. Events with vector bosons are among those that parton shower programs have the greatest difficulty reproducing. This is illustrated in Fig. 29 (left), which shows the 'integrated E_T spectrum for the Nth jet' in events with a Z-boson, i.e., the cross section for the Nth jet to have a transverse energy above E_T. Results are given both from HERWIG and from ALPGEN, which provides an exact LO (tree-level) prediction for each jet's spectrum.
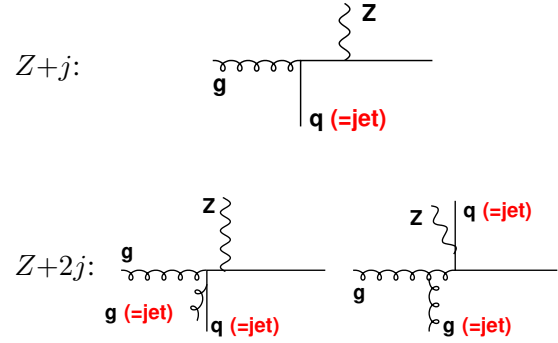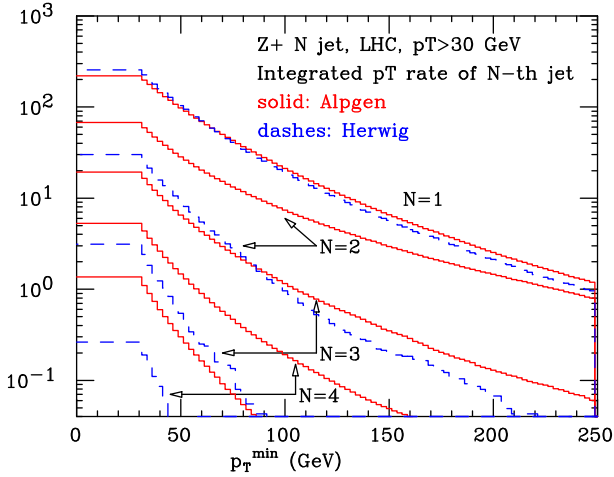
Fig. 29: Left: the cross section for the Nth jet to have a transverse energy above a given E_T, in (14 TeV) LHC events with a Z-boson, as calculated with HERWIG and a tree-level (LO) prediction from ALPGEN. Figure taken from Ref. [46]. Right: kinematic configurations contributing to Z+jet and $Z + 2$ jet events.

The distribution for the first jet is fine: this is by construction, since HERWIG (like PYTHIA) includes the full matrix element for Z+parton production. What is shocking is the result for the second (and higher) jets, for which the E_T spectra are in complete disagreement between HERWIG and ALPGEN.

At first sight it is mysterious how HERWIG could be doing such a good job for pure jet production, Fig. 28, yet such a poor job when there's also a Z-boson in the event, Fig. 29. The diagrams in the right-hand part of Fig. 29 help explain what's going on. HERWIG generates hard configurations like those in the upper line, labelled Z+j. Events with two jets are generated in HERWIG by emission of a gluon off (say) the high-p_t quark. However, there are also events (bottom right) which look like a dijet event with a Z-boson radiated off a quark line. Since at high p_t the Z-boson's mass becomes irrelevant in its production, such diagrams acquire soft and collinear enhancements (just like gluon radiation). However, today's parton-shower Monte Carlos only include QCD showering, not electroweak showering and therefore they are never in a position to start from a dijet event and radiate a Z-boson from it. Therefore they miss a very large part of the cross section.

This example helps illustrate a general feature of the use of Monte Carlos: if you are to trust the results, it is crucial that you know what you have asked the Monte Carlo to generate and whether the observable you are interested in is truly likely to be dominated by what the Monte Carlo can generate.

4.4 Combining fixed-order and parton-shower methods

In the above subsections we saw various strengths and weaknesses of different predictive techniques: NLO codes give predictions with well controlled normalizations, for a reasonable range of processes, as long as one isn't faced with observables that involve disparate scales. Tree-level (LO) predictions can be generated up to quite high multiplicities for a broad range of processes, though without good control of the normalization (i.e., often no better than a factor of two). And parton shower Monte Carlos provide reliable behaviour in soft-collinear regions, giving a fully exclusive final state, though they have normalizations which at best are no better than LO normalizations and sometimes they do dramatically badly in reproducing multi-jet structure.

It is natural to ask whether one can develop tools that combine (or merge) the advantages of all three. This is an active research topic, and here we will just outline the ideas behind two well-established merging tasks: combination of different multiplicity LO matrix elements with parton showers; and combination of NLO and parton shower predictions.
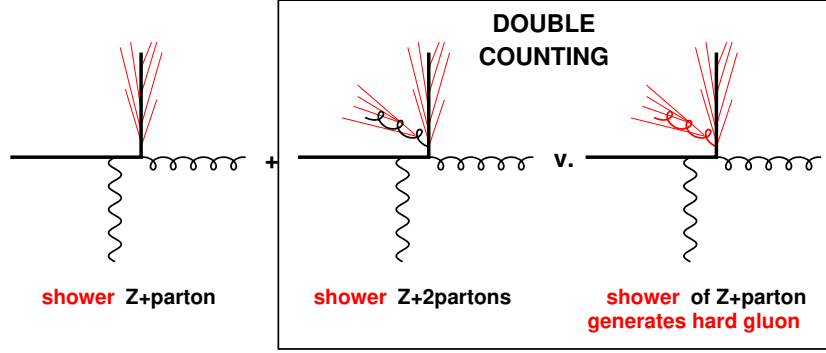
Fig. 30: Illustration of the double-counting issues that can arise if one attempts to shower Z+parton and Z+2-parton events

4.4.1 Matrix elements with parton showers (MEPS)

Suppose you ask for Z+jet production as the initial hard process in PYTHIA or HERWIG. As we saw above, these programs contain the correct matrix element (ME) for Z+parton production, but do a very bad job of estimating Z+2 jet production.

One naive solution to this problem would be to generate Z+2-parton events with ALPGEN, MAD-GRAPH, or some other preferred LO ME tool and then ask HERWIG or PYTHIA to shower those configurations. However, if one showers both Z+parton and Z+2-parton events, then one is faced with a double-counting issue, as illustrated in Fig. 30. In some events (left) the showering off the Z+parton configuration just leads to soft and collinear emissions. Similarly off the Z+2-parton events (middle). However, sometimes (right) the showering off the Z+parton configuration leads to the production of a relatively hard, large-angle gluon that is in the same phase-space region that is already covered in the Z+2-parton event sample.

Two main methods exist to avoid this double counting: CKKW matching [94] and MLM matching [95]. The latter, named after its inventor, M. L. Mangano, is the one we will describe here (it is the simpler of the two). Let's examine the basics of how it proceeds:

– Introduce a (dimensionful) transverse momentum cutoff scale Q_{ME} and a (dimensionless) angular cutoff scale R_{ME} for matrix element generation.
– Generate tree-level events for Z+1-parton, Z+2-partons, ... up to Z+N-partons, where all partons must have $p_t > Q_{ME}$ and be separated from other partons by an angle greater than R_{ME} (we will discuss the definition of this 'angle' later in Section 5). The numbers of events that one generates in the different samples are in proportion to their cross sections with these cuts.
– For each tree-level event from these samples, say one with n partons, shower it with your favourite parton-shower program.
– Apply a jet algorithm to the showered event (choose the algorithm's angular reach R to be $\gtrsim R_{ME}$) and identify all jets with $p_t > Q_{\mathrm{merge}}$, where the merging scale Q_{merge} is to be taken $\gtrsim Q_{ME}$.
– If each jet corresponds to one of the partons (i.e., is nearby in angle) and there are no extra jets above scale Q_{merge}, then accept the event. (For the sample with $n = N$, the condition is that there should be no extra jets with $p_t > p_{tN}$.)
– Otherwise, reject the event.

The action of the MLM procedure on the events of Fig. 30 is illustrated in Fig. 31, showing which events would be accepted and which ones rejected. One immediately sees how the double-counting issue disappears: in the rightmost event, the showering of a Z+parton event leads to an extra jet above Q_{merge}; since this event now has more jets above Q_{merge} than it had partons, it is rejected. In contrast the middle
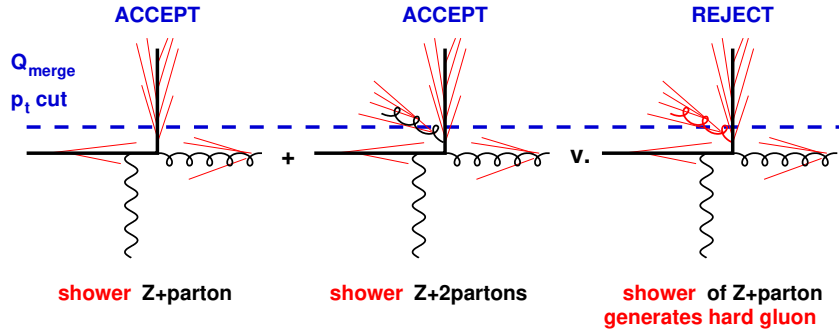
39

Fig. 31: Illustration of the application of the MLM matching procedure to the events of Fig. 30, with the Q_{merge} p_t cutoff represented by the dashed line.

event, which also has two jets above Q_{merge}, was generated from a $Z+2$-parton event and is accepted. So is the leftmost event with only one jet, starting from a $Z+1$-parton event.

By providing a remedy for the double-counting issue, ensuring that the hard jets always come just from the matrix element, the MLM procedure also ensures that hard jets above Q_{merge} have distributions given by the tree-level matrix-element calculations.

The rejection of extra jets also plays another important role: when there are big differences in scales between the jets and Q_{merge} (or between different jets), the Monte Carlo showering would want to 'fill up' that phase space with emissions. However, whenever it does so, the event gets rejected by the matching procedure. As long as the Monte Carlo is carrying out a reasonable showering of these multi-parton events,[8] then the procedure is equivalent to the introduction of a Sudakov form factor encoding the probability of not emitting radiation.

The above 'MLM' merging of matrix-elements and parton showers (MEPS) is the main procedure available with ALPGEN, for use with both HERWIG and PYTHIA. It is also provided (in a variant form) within MADGRAPH. The SHERPA Monte Carlo also has its own matrix-element generator(s) and provides 'CKKW' MEPS matching [94], which instead of the veto steps of MLM matching, uses an analytical calculation of the Sudakov form factors. These and other matrix-element/parton-shower merging schemes are discussed in detail in Ref. [95]. They all share the feature of a matching scale to separate the region under the 'responsibility' of matrix elements and that delegated to the parton shower. In all cases physics predictions well above the matching scale should be independent of the scale. Additionally, distributions at scales around the matching scale should be reasonably smooth, as long as the matching scale has been chosen in a region where the parton shower can be expected to give a reasonable description of emission (for caveats, see Ref. [96]; for a method that avoids the need for a matching scale, see Ref. [97]).

MEPS predictions (as well as other predictive methods) are compared to experimental results for $Z+2$-jet production in Fig. 32 (bottom right). The MEPS results show good agreement for the shape of the observable in question, the p_t distribution of the second jet, and they are much more successful than plain parton-shower predictions. Since their normalizations are based on LO calculations, they do, however, suffer from substantial scale dependence, which is much larger than the scale uncertainty one would obtain at NLO (bottom left).

As a result of their considerable success in describing the shapes of experimental distributions, MEPS predictions have become one of the main tools for a broad range of Tevatron and LHC analyses, especially those involving complex final states.

---

[8]This can depend on subtleties of how the Monte Carlo showers multi-parton events and the communication of information on colour flows between the fixed-order program and the Monte Carlo.
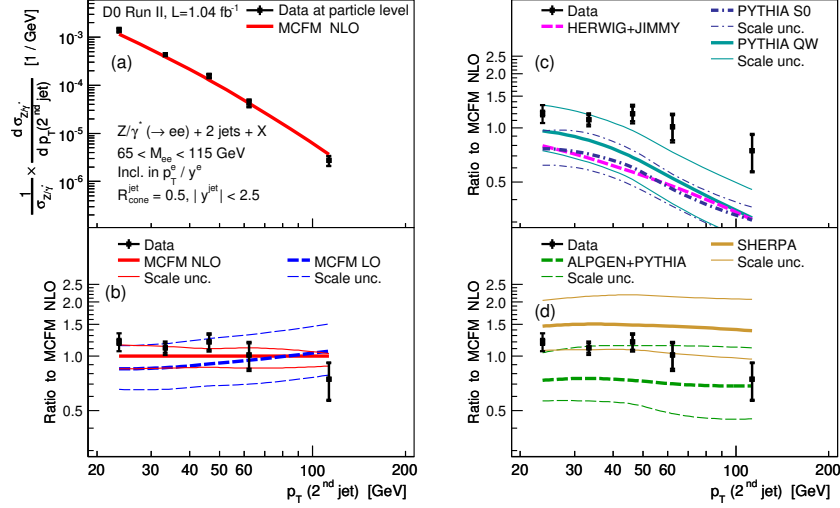
Fig. 32: Cross section for Z+2-jet events at the Tevatron, differential in the transverse momentum of the second jet, as measured by DØ [98]. Comparisons are shown to LO and NLO predictions (from MCFM), parton-shower predictions (PYTHIA, HERWIG) and merged matrix-element + parton-shower predictions (ALPGEN+PYTHIA, SHERPA).

4.4.2 Parton showers and NLO accuracy

We've seen how to obtain a parton shower structure with LO accuracy for a range of different jet multiplicities. However, given the large uncertainty on the normalizations of the predictions, there would be considerable advantages to obtaining NLO accuracy. One might think that since we've had NLO predictions and parton shower predictions for a couple of decades, such a task should not be too hard. The main difficulty comes from the fact that NLO predictions involve divergent event weights (cf. Fig. 24), which aren't even positive definite. Two approaches are in use to get around this problem, the MC@NLO [99] and POWHEG [100] methods.

The idea behind the MC@NLO approach is to 'expand' the Monte Carlo parton shower to first order in $\alpha_{\rm s}$. I.e., the Monte Carlo's parton showers already contain some (partially wrong) estimate of the true NLO corrections and the aim is to figure out what that estimate is. This requires a deep understanding of the Monte Carlo program. As a next step, one calculates the difference between the true NLO contributions and the partial ones included in the Monte Carlo. One of the keys to the MC@NLO method is that as long as the Monte Carlo gives the correct pattern of soft and collinear splitting (which it is supposed to), then the differences between true NLO and Monte Carlo partial NLO should be *finite*. Then one can generate partonic configurations with phase-space distributions proportional to those finite differences and shower them.

Symbolically, if we imagine a problem with one phase-space variable, say energy E, then we can write the 'expansion' of the Monte Carlo cross section as

$$\sigma^{MC} = 1 \times \delta(E) + \alpha_{\rm s}\sigma_{1R}^{MC}(E) + \alpha_{\rm s}\sigma_{1V}^{MC}\delta(E) + \mathcal{O}\left(\alpha_{\rm s}^2\right) \tag{67}$$

where $\sigma_{1R}^{MC}(E)$ is the coefficient of $\alpha_{\rm s}$ for real emission of a gluon with energy E in the Monte Carlo and σ_{1V}^{MC} is the (divergent) coefficient of the virtual corrections. The MC@NLO prediction is then given by

$$\text{MC@NLO} = \text{MC} \times \left(1 + \alpha_{\rm s}(\sigma_{1V} - \sigma_{1V}^{MC}) + \alpha_{\rm s}\int dE(\sigma_{1R}(E) - \sigma_{1R}^{MC}(E))\right), \tag{68}$$

where $\sigma_{1R}(E)$ and σ_{1V} are the true NLO real and virtual coefficients. Each term in (small) brackets in Eq. (68) should separately be finite and corresponds to a given topology of event to be showered: a LO
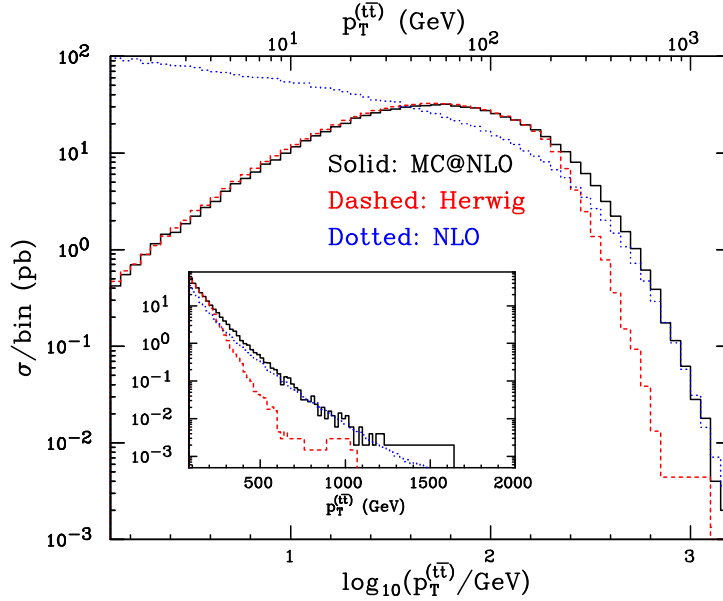
41

Fig. 33: The transverse-momentum distribution of $t\bar{t}$ pairs in the MC@NLO approach, compared to the plain HERWIG result (rescaled by the $\sigma_{\mathrm{NLO}}/\sigma_{\mathrm{LO}}$ K-factor) and to the NLO calculation. Shown for a $14\,\mathrm{TeV}$ LHC. Figure taken from Ref. [105].

topology for the "1" and $(\sigma_{1V} - \sigma_{1V}^{MC})$ terms, and a NLO real topology for the $(\sigma_{1R}(E) - \sigma_{1R}^{MC}(E))$ term. (A more complete and clearer 'toy' explanation is given in the original MC@NLO paper [99], which makes for very instructive reading.)

The MC@NLO approach has the practical characteristic that all event weights are ± 1. Quite a range of processes are available in the MC@NLO approach for HERWIG, including the production of a Higgs boson, single and double vector-bosons, a heavy-quark pair, various single-top processes, and $H + W$ and $H + Z$. The characteristic of these processes is that they are nearly all free of light jets at LO (except for one of the single-top processes), because this simplifies the set of divergences that need to be dealt with. Very recently one PYTHIA [101] and a couple of HERWIG++ processes (Ref. [102] and references therein) were also interfaced within the MC@NLO approach.[9]

An alternative to MC@NLO is POWHEG [100]. It aims to avoid the (small fraction of) negative weights that are present in MC@NLO and also seeks to be less tightly interconnected with a specific Monte Carlo program. The principle of the POWHEG approach is to write a simplified Monte Carlo that generates just one emission beyond LO. This single-emission Monte Carlo is designed in such a way as to be sufficient to give the correct NLO result. It essentially works by introducing a Sudakov form factor such as Eq. (65) in which the contents of the square bracket are replaced by the integral over the exact real radiation probability above k_t (plus a constant term that accounts for the finite part of the 1-loop correction). Then emissions with transverse momenta below the scale of the first emission are left to the default Monte Carlo program, for example HERWIG or PYTHIA (implementing a transverse momentum veto to ensure that nothing is generated above the scale of the first, POWHEG, emission). The range of processes available within POWHEG is gradually catching up with that for MC@NLO, and it is hoped that this will be helped by the availability of systematic tools to help with the development of new processes [104].

An illustration of a result with MC@NLO is given in Fig. 33 for the transverse-momentum distribution of a $t\bar{t}$ pair. It illustrates how MC@NLO reproduces the HERWIG shape in the low-p_t region

---

[9]In the time since the original version of these lectures were written up, many more processes have been implemented for HERWIG++ [103].
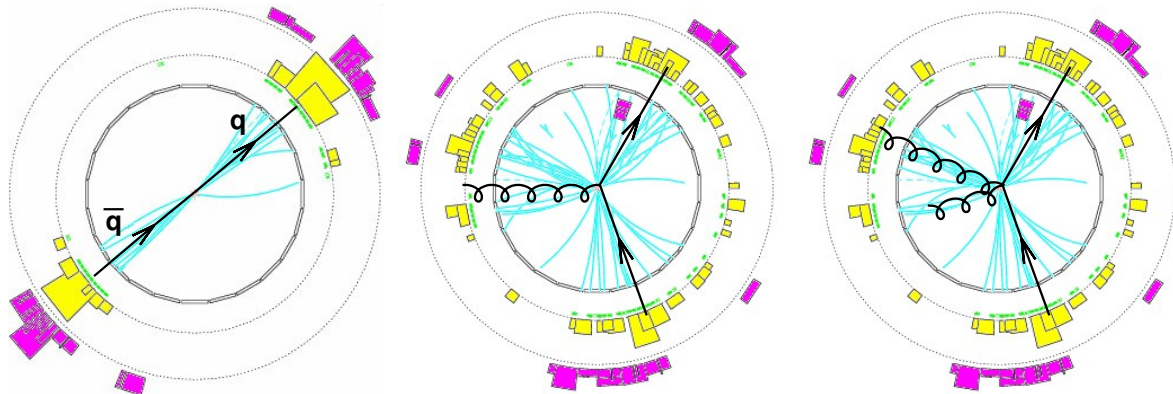
Fig. 34: Left: an e^+e^- event that can be interpreted as having a 2-jet, $q\bar{q}$-like structure; middle: an event that can be interpreted as having a 3-jet, $q\bar{q}g$, structure; right: the same event reinterpreted as having a 4-jet structure, $q\bar{q}gg$.

(with NLO normalization), the NLO distribution at high p_t, and that neither HERWIG nor plain NLO are able to properly describe both regions.

4.5 Summary

In this section, we have seen quite a range of different predictive methods for QCD at hadron colliders. Two big classes of predictive methods exist: partonic fixed-order calculations, which have well controlled accuracy, but wildly fluctuating positive and negative event weights; and Monte Carlo parton shower tools, which give a much more complete description of events (and uniform event weights).

Development is active on both sets of tools individually. On one hand we've mentioned the challenge of having a broader range of processes known at NLO and a handful at NNLO. And though we have not really touched on it, there is also a very active programme to develop parton shower Monte Carlos in C++ as replacements for venerable but ageing Fortran codes like PYTHIA 6.4 and HERWIG 6.5.

In addition there are methods with the advantages of both fixed-order and parton-shower programs. It is widespread nowadays to merge different LO 'tree-level' predictions (e.g., Z+parton, Z+2-partons, Z+3-partons, etc.) together with parton showers. And for simple processes, those with at most one light parton in the final state, it is possible to combine NLO accuracy with the full parton-shower description. Ultimately the hope is to be able to combine Z+jet, Z+2-jets, Z+3-jets all at NLO accuracy also merging them with parton showers, so as to obtain accurate descriptions for the whole range of processes that are relevant at the LHC, both as backgrounds and as signals of new particles and new physics.

5 Jets

The concept of a jet has already arisen in various contexts, so in this final section we will examine jet-related ideas in more detail.

Consider the three events of Fig. 34. In the left-hand one, one interpretation is that we're seeing an $e^+e^- \to q\bar{q}$ event, in which there has been soft and collinear showering followed by a transition to hadrons. This is a classic picture of a '2-jet' event. The middle event is more complex: energy flow is not limited to two cones. One interpretation of the event is that a $q\bar{q}$ pair has emitted a hard gluon g, and all three have undergone soft and collinear showering. However, the same event can also be interpreted (right) as a $q\bar{q}gg$ event, with further soft and collinear showering. Deciding between these two interpretations means choosing just how hard and separated in angle an emission has to be in order for it to be considered a separate jet (cf. the angular and energy parameters, δ and ϵ, in our discussion of the 2-jet cross section in Section 2.3.2).

In making this choice, it would be highly painful to visually inspect each of the $\mathcal{O}\left(10^9\right)$ events
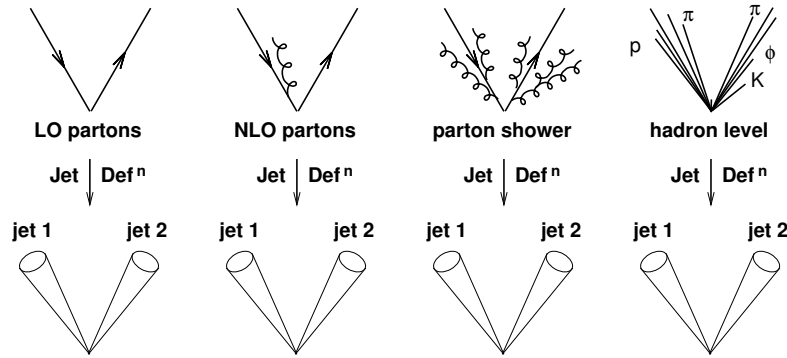
Fig. 35: The application of a jet definition to a variety of events that differ just through soft/collinear branching (and hadronization), should give identical jets in all cases.

written to disk every year at the LHC. Instead one uses a set of rules, a 'jet definition', by which a computer can take a list of particle momenta for an event (be they quark and gluons, or hadrons, or even calorimeter deposits), and return a list of jets. If one modifies an event just through soft and collinear emission, then the set of jets should not change, i.e., the result of applying the jet definition should be insensitive to the most common effects of showering and hadronization, as illustrated in Fig. 35.

Jets are central to collider physics: both theory and experimental results are often presented in terms of jet cross sections, and thus jets provide the meeting point between the two. As we saw in Section 4.4.1, jets are also used to assemble together different kinds of theory predictions. And jets are an input to almost all physics analyses: to new physics searches (since new particles may decay to quarks or gluons, giving jets), in Higgs searches, top physics, Monte Carlo validation, fits of PDFs, etc.

5.1 Jet definitions

The construction of a jet involves different considerations:

– Which particles are grouped together into a common jet? The set of rules that one follows for deciding this is usually known as a jet algorithm, and it comes with parameters that govern its exact behaviour. A common parameter is R which determines the angular reach of the jet algorithm.
– How do you combine the momenta of particles inside a jet? One needs to specify a 'recombination scheme'. The most common is to simply add the 4-vectors of the particles (the 'E-scheme'). This gives jets that are massive (so jets cannot be thought of as a direct stand-in for partons, which are massless).

Taken together, the algorithm, its parameters and the recombination scheme specify a 'jet definition'.

Two broad classes of jet definition are in common use: cone algorithms, which take a top-down approach, and sequential recombination algorithms, based on a bottom-up approach. Below we'll give a brief discussion of each kind of algorithm, referring the reader to Ref. [106] for a more complete description of all the variants that exist.

5.2 Cone algorithms

There are many types of cone algorithm, but all rely on the idea that soft and collinear branching doesn't modify the basic direction of energy flow.

One of the simpler cone algorithms (we'll call it IC-PR, for iterative cone with progressive removal of particles) is that used by the CMS experiment during much of their preparation for LHC running. One first sorts all particles according to their transverse momentum, and identifies the one with largest transverse momentum. This is referred to as a seed particle, s. One draws a cone of radius R around the

Collinear Safe		**Collinear Unsafe**	

$$\alpha_s^n \times (-\infty) \qquad \alpha_s^n \times (+\infty) \qquad \alpha_s^n \times (-\infty) \qquad \alpha_s^n \times (+\infty)$$

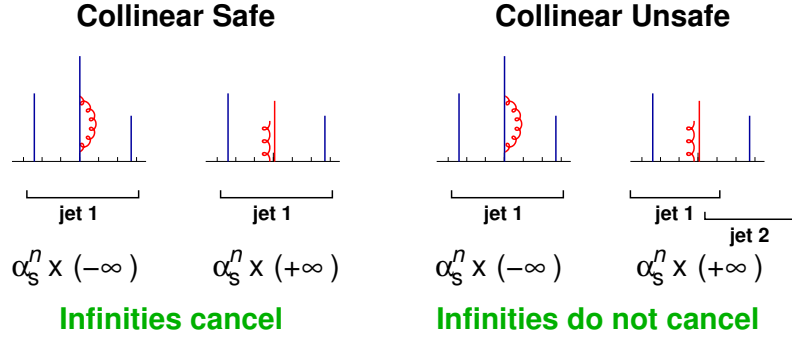**Infinities cancel**   **Infinities do not cancel**

Fig. 36: Sample events illustrating the result of applying collinear safe and unsafe jet algorithms. The height of a given line corresponds to the particle's transverse momentum, its horizontal position to its rapidity ($\phi = 0$ for all particles here). Left: expectations for the behaviour of a collinear-safe jet algorithm, where the jets found by the algorithm should be independent on the collinear splitting of the hardest particles. Right: in a collinear unsafe algorithm such as the IC-PR type CMS cone, the splitting of the central hard particle causes the leftmost particle to become the hardest in the event, leading to a two-jet rather than a one-jet event.

seed — in hadron-collider variables this means identifying all particles with $\Delta R_{is}^2 = (y_i - y_s)^2 + (\phi_i - \phi_s)^2 < R^2$, where $y_i = \frac{1}{2} \ln \frac{E_i + p_{zi}}{E_i + p_{zi}}$ is the rapidity of particle i, ϕ_i its azimuth, and y_s and ϕ_s the rapidity and azimuth of the seed. One then identifies the direction of the sum of the momenta of those particles. If it doesn't coincide with the seed direction then one uses that sum as a new seed direction, and iterates until the sum of the cone contents coincides with the previous seed. This is referred to as finding a stable cone. The particles inside that stable cone make a jet, and they're removed from the list of particles in the event. The procedure is then repeated on the particles that remain, removing particles each time one finds a stable cone ($\rightarrow$ jet), until no particles remain and one has the complete set of jets. Of these jets one retains only those above some transverse-momentum threshold $p_{t,\min}$.

There is one major drawback to the above procedure: the use of the particles' p_t's to decide which one to take as the first seed. This is problematic, because particle p_t's are not collinear safe quantities. As illustrated in the two right-hand events of Fig. 36, in an IC-PR algorithm, if the hardest particle undergoes a collinear splitting then this can result in another particle in the event becoming the 'new' hardest particle, giving a different set of final jets as compared to events without the splitting. Thus in the example of Fig. 36 there is a divergent (real, positive) contribution to the 2-jet cross section and a separate divergent (1-loop virtual, negative) contribution to the 1-jet cross section. In contrast, for a collinear-safe algorithm (two leftmost events), the collinear-splitting of the hardest particle does not change the set of final jets. Then the real and virtual divergences both contribute to the 1-jet cross section and so cancel each other.

Collinear unsafety means that certain cross sections cannot be calculated at NLO (or sometimes NNLO) — one will only obtain nonsensical infinite answers. Furthermore, even if one is working at some low perturbative order which is not divergent (e.g., LO), the fact that higher orders diverge means that the convergence of the whole perturbative series becomes questionable, compromising the usefulness even of the low orders.

Over the past two decades there has been significant discussion of such problems. There are many other variants of cone algorithm, and nearly all suffer from problems either of collinear safety, or infrared safety. One class that has been widely used at the Tevatron avoids the ordering of initial seeds, and instead obtains stable cones using all possible particles as seeds: stable cones are not immediately converted into jets, but instead, once one has the list of the stable cones found by iterating from all possible seeds one then uses a 'split–merge' procedure to decide how particles that appear in multiple stable cones should be unambiguously assigned to a single jet.
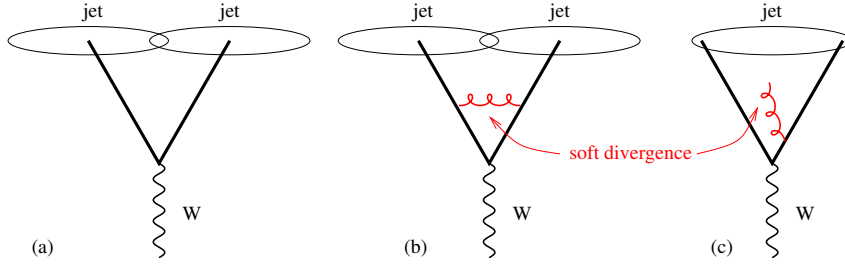
Fig. 37: Configurations illustrating IR unsafety of iterative cone algorithms with a split–merge procedure, in events with a W and two hard partons. The addition of a soft gluon converts the event from having two jets to just one jet. In contrast to Fig. 36, here the explicit angular structure is shown (rather than p_t as a function of rapidity).

This procedure avoids the collinear-safety issue (the order of particles' p_t's no longer matters), however, it turns out that it instead introduces an infrared-safety issue: adding an extra soft particle creates a new seed, which can lead to an extra stable cone being found, which feeds through the split–merge procedure, altering the final set of (hard) jets. This tends to happen when two hard particles are separated by a distance ΔR that satisfies $R < \Delta R < 2R$ (so that a cone centred on either fails to capture the other and each hard particle leads to its own jet) and one adds a soft particle in between the two (so that a cone centred on the soft particle includes both hard ones, which then end up in a single jet), as illustrated in Fig. 37. A partial fix for the problem was given in Ref. [107] (adopted by the Tevatron in Ref. [108]), which involves adding extra seeds at the midpoints between all pairs of stable cones and uses those as new starting points for finding additional stable cones before starting the split–merge step. A full fix involves a non-seed-based approach to exhaustively finding all possible stable cones in an event, in an algorithm known as the Seedless Infrared Safe Cone (SISCone) [109].

5.3 Sequential-recombination algorithms

Sequential-recombination jet algorithms take a bottom-up approach to constructing jets, as if they were inverting the sequence of splittings of the parton shower. Of course that sequence doesn't really exist unambiguously, since gluon emission is a quantum-mechanical process involving coherent emission from all colour sources in an event. However, for collinear emissions the picture that there is a single 'emitter' is not a poor approximation.

5.3.1 The e^+e^- k_t algorithm

The most widely used sequential recombination algorithm to date is the k_t algorithm, originally formulated for e^+e^- events [110]. Recall, from Eq. (26), that the soft and collinear limit of the gluon-emission probability for $a \to ij$ is

$$dS \simeq \frac{2\alpha_{\mathrm{s}} C_i}{\pi} \frac{dE_i}{\min(E_i, E_j)} \frac{d\theta_{ij}}{\theta_{ij}} \, , \qquad (69)$$

where C_i is C_A (C_F) if a is a gluon (quark), and where we've written $\min(E_i, E_j)$ in the denominator to avoid specifying which of i and j is the soft particle.

The essential idea of the k_t algorithm is to define a 'distance measure' y_{ij} between every pair of particles i, j,

$$y_{ij} = \frac{2\min(E_i^2, E_j^2)(1 - \cos\theta)}{Q^2} \, . \qquad (70)$$

In the collinear limit, y_{ij} reduces to $\min(E_i^2, E_j^2)\theta_{ij}^2$, which is the relative transverse momentum between particles i and j (hence the name k_t algorithm), normalized to the total visible (or sometimes centre-of-mass) energy Q. Apart from the normalization, this is just what appears in the denominator of the
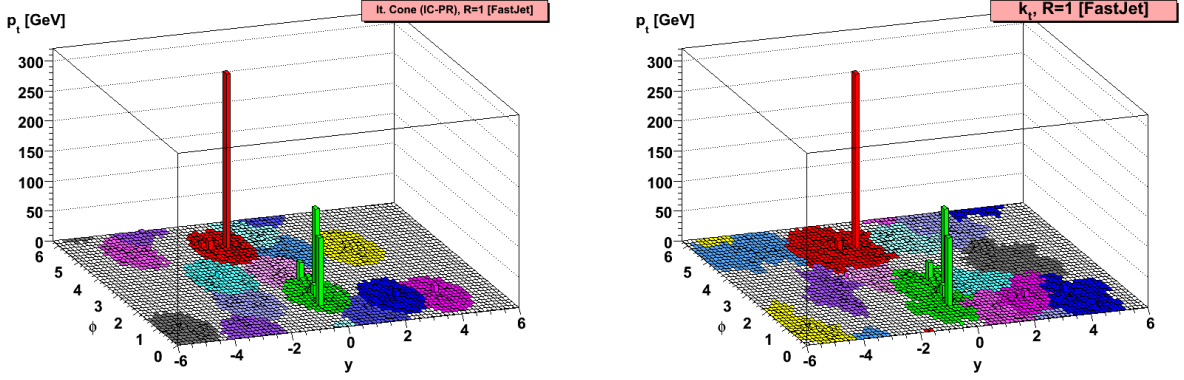
Fig. 38: Regions of the y–ϕ plane covered by jets in an illustrative (simulated) hadron-collider event with an IC-PR type cone algorithm (left) and the inclusive longitudinally-invariant k_t algorithm (right). The jet finding was carried out with the FASTJET package [113,114], with the addition of very soft ghost particles to trace the extent of the jets [115].

splitting probability, Eq. (69), so that pairs of particles that would arise from a splitting with a strong divergence are considered very close together.

The algorithm works by identifying the pair of particles that has the smallest y_{ij} and recombining them into a single particle (also called a 'pseudojet'). It then recalculates all y_{ij} distances taking into account the new particle, and again recombines the pair that's closest. The procedure is repeated until all y_{ij} distances are above some threshold y_{cut}, at which point the pseudojets that remain become the event's jets.

5.3.2 The k_t algorithm for hadron collisions

For hadron collisions and deep-inelastic scattering, the version of the k_t algorithm that is most commonly used reads as follows [111,112]. For every pair of particles define a (dimensionful) inter-particle distance d_{ij},

$$d_{ij} = \min(p_{ti}^2, p_{tj}^2)\frac{\Delta R_{ij}^2}{R^2}\,, \tag{71}$$

where R is a parameter whose role is similar to that of R in cone algorithms. Also define a beam distance for every particle,

$$d_{iB} = p_{ti}^2\,. \tag{72}$$

The algorithm proceeds by searching for the smallest of the d_{ij} and the d_{iB}. If it is a d_{ij} then particles i and j are recombined into a single new particle. If it is a d_{iB} then i is removed from the list of particles, and called a jet. This is repeated until no particles remain.

Note that the distance in Eq. (71) just reduces to that of Eq. (70) in the collinear limit (modulo Q^2 normalization). So one is still dealing with the relative transverse momentum between pairs of particles. As with cone algorithms, in this 'inclusive longitudinally invariant k_t algorithm,' arbitrarily soft particles can form jets. It is therefore standard to place a $p_{t,\min}$ cutoff on the jets one uses for 'hard' physics.

One can verify that R in the k_t algorithm plays a similar role to R in cone algorithms, using the following observations: if two particles i and j are within R of each other, i.e., $\Delta R_{ij} < R$, then $d_{ij} < d_{iB}, d_{jB}$ and so i and j will prefer to recombine rather than forming separate jets. If a particle i is separated by more than R from all other particles in the event then it will have $d_{iB} < d_{ij}$ for all j and so it will form a jet on its own.

Despite this similarity to the behaviour of cone algorithms for pairs of particles, the k_t algorithm gives jets that 'look' somewhat different. Figure 38 illustrates what happens when one clusters a simulated hadron-collider event with an IC-PR type cone algorithm and with the k_t algorithm. In both cases the figure shows (in a given colour) the calorimeter cells that are included in each jet. For the IC-PR algorithm the hardest jets all appear circular, as expected given the use of cones in the definition of the algorithm (in cone algorithms with split–merge steps, the jets are often not circular, because of the less trivial mapping from stable cones to jets). For the k_t algorithm, the jets have quite irregular (or jagged) edges, because many of the soft particles cluster together early in the recombination sequence (owing to their small p_t and hence d_{ij}) in patterns that are determined by the random distributions of those particles in p_t and rapidity and azimuth.

The irregularity of the jet boundaries has often been held against the k_t algorithm. One reason is that it makes it harder to calculate acceptance corrections: for example, if you know that some part of a detector is misbehaving, it is not obvious how far a k_t jet must be from that part of the detector in order not to be affected by it. Another reason relates to the linearity of the dependence of the jet momentum on soft-particle momenta: in the IC-PR algorithm, the hard core of the jet essentially determines the jet boundary, and the algorithm depends linearly on the momenta of any soft particles within the boundary, and is independent of particles outside it. In the k_t algorithm, varying the momentum of one soft particle in the vicinity of the jet core can affect whether it and other soft particles get clustered with that core or not. This can complicate energy calibrations for the jet algorithm, though techniques exist to correct for this to some extent (jet-area-based subtraction [116], which leaves just a residual term known as back-reaction [115]).

A feature of the k_t algorithm that is attractive is that it not only gives you jets, but also assigns a clustering sequence to the particles within the jet. One can therefore undo the clustering and look inside the jet. This has been exploited in a range of QCD studies (e.g., Ref. [117]), and also in discussions of searches of hadronic decays of boosted massive particles such as W, H, or Z bosons, top quarks, or new particles (early examples include Refs. [118, 119]; for more recent examples, see the reviews in Refs. [106, 120]). Jet substructure studies are also often carried out with the Cambridge/Aachen (C/A) algorithm [121, 122], which essentially replaces $p_{ti} \to 1, p_{tj} \to 1$ in Eqs. (71,72) but is otherwise like the k_t algorithm.

5.3.3   The anti-k_t algorithm

It turns out that it is possible to design a sequential-recombination algorithm with many of the nice properties of cone algorithms via a simple modification of the k_t algorithm's distance measures [123]:

$$d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2} \,, \tag{73a}$$

$$d_{iB} = \frac{1}{p_{ti}^2} \,. \tag{73b}$$

The correspondence with the divergences of Eq. (69) is partially lost: objects that are close in angle still prefer to cluster early, but that clustering tends to occur with a hard particle (rather than necessarily involving soft particles). This means that jets 'grow' in concentric circles out from a hard core, until they reach a radius R, giving circular jets just as with the IC-PR cone, as shown in Fig. 39. However, unlike the IC-PR cone, this 'anti-k_t' algorithm is collinear (and infrared) safe, meaning that it is safe to use with fixed-order QCD predictions. This, combined with the fact that it has been implemented efficiently in the FASTJET jet-finding code [113, 114], has led to it being adopted as the default jet algorithm by both the ATLAS and CMS collaborations.

Note that the anti-k_t algorithm does not provide useful information on jet substructure: if a jet contains two hard cores, then the k_t (or C/A) algorithms first reconstruct those hard cores and merge the
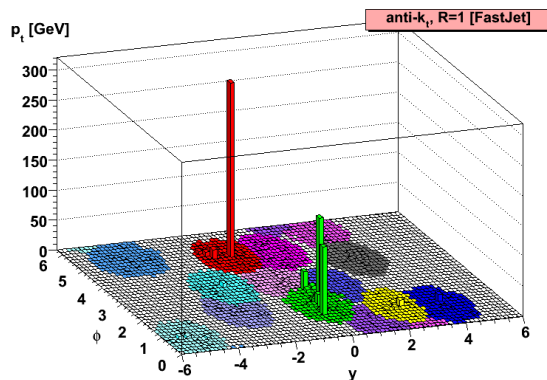
Fig. 39: As in Fig. 38, but shown for the anti-k_t algorithm

resulting two subjets. The anti-k_t algorithm will often first cluster the harder of the two cores and then gradually agglomerate the contents of the second hard core.

5.4 Using jets

Because of the intricacies of calibrating jets, past experiments have tended to concentrate their efforts on just one or two jet definitions, which are used across the board for QCD studies, top-quark studies, Higgs and new physics searches. Typically the choices have been for cone-type algorithms with R values in the range 0.4–0.7.

At the LHC there are prospects of experimental methods (for example topoclusters at ATLAS and particle flow at ATLAS) that make it easier to use a broad range of jet definitions. We've already mentioned, above, the use of jet substructure in new physics searches. These methods will definitely benefit from experiments' flexibility in choosing different jet definitions (for example, many of the proposed searches use quite a large jet radius, $R \sim$ 1–1.5). More generally, when using jets, it is important to establish what you are trying to get out of the jet algorithm, what it is that you're trying to optimize.

Different studies will want to optimize different things. For example in QCD studies, such as the inclusive jet-spectrum measurement that goes into PDF fits, one criterion might be to choose a jet definition that minimizes the non-perturbative corrections that need to be applied when comparing perturbative predictions with data.

In searches for new physics, for example when looking for a resonance in the dijet invariant mass spectrum, the criterion might be the following: given a narrow resonance, what jet definition will lead to the narrowest peak in the dijet mass spectrum? It turns out that the answer depends significantly on properties of the resonance you're trying to reconstruct. Figure 40 illustrates how SISCone with $R = 0.5$ does well for reconstructing a $100\,\text{GeV}$ resonance that decays to $q\bar{q}$. While for a $2\,\text{TeV}$ resonance decaying to gg, you're better off with an algorithm (such as SISCone, or others supplemented with the trick of filtering and/or variants [126, 127]) with a substantially larger radius, $R \sim 1$ [124, 125, 128, 129]. And if you find a resonance, you might then want to know which jet definition will allow you to measure its mass most accurately, which may not be the same choice that allowed you to find the resonance most easily.

Understanding analytically which jet choices work well is the subject of ongoing theoretical work [115, 130, 131], which involves figuring out how perturbative radiation, hadronization, the underlying event and pileup all come together to affect jet momenta.
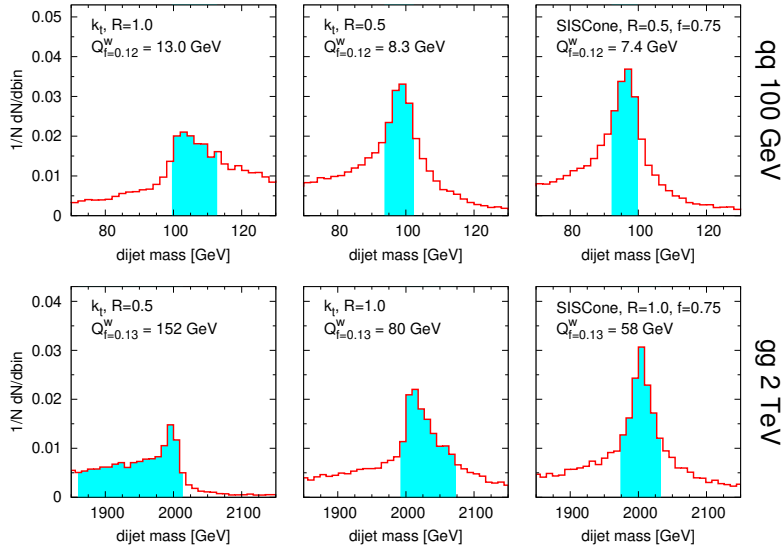
Fig. 40: Illustration of the impact of different jet definition on the quality of reconstruction of dijet invariant mass peaks for a $100\,\mathrm{GeV}$ resonance decaying to $q\bar{q}$ (top) and a $2\,\mathrm{TeV}$ resonance decaying to gg (bottom). The $Q^w_{f=0.12}$ indicates the width of the region that contains 12% of the resonance events (corresponding to about 25% of the resonance events that pass basic selection cuts). Smaller numbers imply better reconstruction. Figure taken from Ref. [124]. Many further plots available from Ref. [125].

6 Conclusions

In these lectures we have covered material that ranges from the QCD Lagrangian through to e^+e^- scattering, soft and collinear divergences, PDFs, fixed-order calculations, parton showers and jets. It is my hope that the material here will have given you enough background information to at least understand the basics of these different concepts.

The aspects of QCD that we have examined will play a role in nearly all analyses at the LHC. Whether because Monte Carlo parton-shower programs are used for estimating detector effects on almost every measurement; or because in searches for new physics one might look for an excess with respect to QCD predictions of backgrounds; or perhaps because an understanding of QCD will make it possible to more clearly pull out a signal of new physics that involves QCD final states.

If you want to find out more about the topics discussed here (or those left out due to time constraints, notably heavy quarks), good places to look include the textbooks [1–3] mentioned at the beginning of these lectures, the references given throughout, and transparencies from the various summer schools dedicated to QCD, for example, the CTEQ (and MCNET) schools [132, 133].

References

[1] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and collider physics*, *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **8** (1996) 1, doi:10.1017/CBO9780511628788.

[2] G. Dissertori, I. G. Knowles and M. Schmelling, *Quantum Chromodynamics: High-Energy Experiments and Theory*, (Clarendon, Oxford, 2002), doi:10.1093/acprof:oso/9780199566419.001.0001.

[3] R. Brock *et al.* [CTEQ Collaboration], *Rev. Mod. Phys.* **67** (1995) 157, doi:10.1103/RevModPhys.67.157 ; see also http://www.phys.psu.edu/~cteq/handbook/v1.1/handbook.pdf.

[4] J. M. Maldacena, *Adv. Theor. Math. Phys.* **2** (1998) 231, doi:10.1023/A:1026654312961,

Int. J. Theor. Phys. **38** (1999) 1113, doi:10.4310/ATMP.1998.v2.n2.a1, arXiv:hep-th/9711200.

[5] S. S. Gubser, I. R. Klebanov and A. M. Polyakov, *Phys. Lett.* **B428** (1998) 105, doi:10.1016/S0370-2693(98)00377-3, arXiv:hep-th/9802109.

[6] E. Witten, *Adv. Theor. Math. Phys.* **2** (1998) 253, doi:10.4310/ATMP.1998.v2.n2.a2, arXiv:hep-th/9802150.

[7] J. Erlich, E. Katz, D. T. Son and M. A. Stephanov, *Phys. Rev. Lett.* **95** (2005) 261602, doi:10.1103/PhysRevLett.95.261602, arXiv:hep-ph/0501128.

[8] L. Da Rold and A. Pomarol, *Nucl. Phys.* **B721** (2005) 79, doi:10.1016/j.nuclphysb.2005.05.009, arXiv:hep-ph/0501218.

[9] S. Durr *et al.*, *Science* **322** (2008) 1224, doi: 10.1126/science.1163233, arXiv:0906.3599 [hep-lat].

[10] J. Smit, *Introduction to quantum fields on a lattice: A robust mate*, *Cambridge Lect. Notes Phys.* **15** (2002) 1, doi:10.1017/CBO9780511583971.

[11] T. DeGrand and C. E. Detar, *Lattice Methods for Quantum Chromodynamics*, (World Scientific, Singapore, 2006), doi:10.1142/6065.

[12] Perspectives in Lattice QCD, Proceedings of the Workshop, Nara International Seminar House, Nara, Japan, 2005, Ed. Y. Kuramashi, (World Scientific, Singapore, 2007), doi:10.1142/6216.

[13] T. Onogi, arXiv:0906.2344 [hep-ph].

[14] M. E. Peskin and D. V. Schroeder, An Introduction To Quantum Field Theory (Addison-Wesley, Reading, MA, 1995).

[15] V. N. Gribov, Lectures at the 12th Winter School of the Leningrad Nuclear Physics Institute, SLAC-TRANS-176 (1977), https://www.slac.stanford.edu/pubs/slactrans/trans01/slac-trans-0176.pdf; S. D. Drell, *in* a Festschrift for Maurice Goldhaber, *Trans. NY Acad. Sci., Series II* **40** (1980) 76, doi:10.1111/j.2164-0947.1980.tb03001.x.

[16] N. K. Nielsen, *Am. J. Phys.* **49** (1981) 1171, doi:10.1119/1.12565; R. J. Hughes, *Phys. Lett.* **B97** (1980) 246, doi:10.1016/0370-2693(80)90593-6; *idem Nucl. Phys.* **B186** (1981) 376, doi:10.1016/0550-3213(81)90076-6; K. Johnson, *in* Asymptotic Realms of Physics, Eds. A. Guth, K. Huang and R. L. Jaffe (MIT Press, Cambridge, MA, 1983).

[17] S. Bethke, *Eur. Phys. J.* **C64** (2009) 689, doi:10.1140/epjc/s10052-009-1173-1, arXiv:0908.1135 [hep-ph].

[18] S. G. Gorishnii, A. L. Kataev and S. A. Larin, *Phys. Lett.* **B259** (1991) 144, doi:10.1016/0370-2693(91)90149-K.

[19] L. R. Surguladze and M. A. Samuel, *Phys. Rev. Lett.* **66**, (1991) 560, doi:10.1103/PhysRevLett.66.560, [Erratum-ibid. **66**, 2416 (1991)].

[20] T. Hebbeker, M. Martinez, G. Passarino and G. Quast, *Phys. Lett.* **B331** (1994) 165, doi:10.1016/0370-2693(94)90958-X.

[21] K. G. Chetyrkin, J. H. Kuhn and A. Kwiatkowski, *Phys. Rep.* **277** (1996) 189, doi:10.1016/S0370-1573(96)00012-9.

[22] P. A. Baikov, K. G. Chetyrkin and J. H. Kuhn, *Phys. Rev. Lett.* **101** (2008) 012002, doi:10.1103/PhysRevLett.101.012002, arXiv:0801.1821 [hep-ph].

[23] P. A. Baikov, K. G. Chetyrkin and J. H. Kuhn, *Nucl. Phys. B Proc. Suppl.* **189** (2009), 49, doi:10.1016/j.nuclphysbps.2009.03.010, arXiv:0906.2987 [hep-ph].

[24] V. S. Fadin, *Yad. Fiz.* **37** (1983) 408; B. I. Ermolaev and V. S. Fadin, *JETP Lett.* **33** (1981) 269 [Pisma Zh. Eksp. Teor. Fiz. **33** (1981) 285], http://www.jetpletters.ac.ru/ps/1505/article_23008.shtml; A. H. Mueller, *Phys. Lett.* **B104**, (1981) 161, doi:10.1016/0370-2693(81)90581-5;

Y. L. Dokshitzer, V. S. Fadin and V. A. Khoze, *Z. Phys.* **C15** (1982) 325, doi:10.1007/BF01614423;
A. Bassetto, M. Ciafaloni and G. Marchesini, *Phys. Rept.* **100** (1983) 201, doi:10.1016/0370-1573(83)90083-2.

[25] Y. L. Dokshitzer, V. A. Khoze, A. H. Mueller and S. I. Troyan, *Basics of perturbative QCD*, (Editions Frontières, Gif-sur-Yvette, 1991), http://www.lpthe.jussieu.fr/~yuri/BPQCD/cover.html.

[26] A. H. Mueller, *Nucl. Phys.* **B213** (1983) 85, doi:10.1016/0550-3213(83)90176-1.

[27] G. Sterman and S. Weinberg, *Phys. Rev. Lett.* **39** (1977) 1436, doi:10.1103/PhysRevLett.39.1436.

[28] J. M. Zanotti, *PoS* **LC2008** (2008) 051, doi:10.22323/1.061.0051.

[29] M. Arneodo *et al.* [New Muon Collaboration], *Nucl. Phys.* **B483** (1997) 3, doi:10.1016/S0550-3213(96)00538-X, arXiv:hep-ph/9610231.

[30] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky and W. K. Tung, *JHEP* **07** (2002) 012, doi:10.1088/1126-6708/2002/07/012, arXiv:hep-ph/0201195.

[31] S. Chekanov *et al.* [ZEUS Collaboration], *Eur. Phys. J.* **C21** (2001) 443, doi:10.1007/s100520100749 arXiv:hep-ex/0105090.

[32] W. K. Tung, H. L. Lai, A. Belyaev, J. Pumplin, D. Stump and C. P. Yuan, *JHEP* **02** (2007) 053, doi:10.1088/1126-6708/2007/02/053, arXiv:hep-ph/0611254.

[33] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, *Phys. Lett.* **B604** (2004) 61, doi:10.1016/j.physletb.2004.10.040, arXiv:hep-ph/0410230.

[34] V. M. Abazov *et al.* [D0 Collaboration], *Phys. Rev. Lett.* **101** (2008) 062001, doi:10.1103/PhysRevLett.101.062001, arXiv:0802.2400 [hep-ex].

[35] A. Abulencia *et al.* [CDF - Run II Collaboration], *Phys. Rev.* **D75** (2007) 092006, doi:10.1103/PhysRevD.75.092006, [Erratum-*ibid.* **D75** (2007) 119901], arXiv:hep-ex/0701051.

[36] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Eur. Phys. J.* **C64** (2009) 653, doi:10.1140/epjc/s10052-009-1164-2, arXiv:0905.3531 [hep-ph].

[37] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Eur. Phys. J.* **C63** (2009) 189, doi:10.1140/epjc/s10052-009-1072-5, arXiv:0901.0002 [hep-ph].

[38] P. M. Nadolsky *et al.*, *Phys. Rev.* **D78** (2008) 013004, doi:10.1103/PhysRevD.78.013004, arXiv:0802.0007 [hep-ph].

[39] R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, J. Rojo and M. Ubiali, *Nucl. Phys.* **B838** (2010) 136, doi:10.1016/j.nuclphysb.2010.05.008, arXiv:1002.4407 [hep-ph].

[40] G. P. Salam and J. Rojo, *Comput. Phys. Commun.* **180** (2009) 120, doi:10.1016/j.cpc.2008.08.010, arXiv:0804.3755 [hep-ph].

[41] W. Furmanski and R. Petronzio, *Phys. Lett.* **B97** (1980) 437, doi:10.1016/0370-2693(80)90636-X.

[42] G. Curci, W. Furmanski and R. Petronzio, *Nucl. Phys.* **B175** (1980) 27, doi:10.1016/0550-3213(80)90003-6.

[43] A. Vogt, S. Moch and J. A. M. Vermaseren, *Nucl. Phys.* **B691** (2004) 129, doi:10.1016/j.nuclphysb.2004.04.024, arXiv:hep-ph/0404111.

[44] S. Moch, J. A. M. Vermaseren and A. Vogt, *Nucl. Phys.* **B688** (2004) 101, doi:10.1016/j.nuclphysb.2004.03.030, arXiv:hep-ph/0403192.

[45] LHAPDF, *LHAPDF Documentation*, http://projects.hepforge.org/lhapdf/.

[46] M. L. Mangano, *Eur. Phys. J.* **C59** (2009) 373, doi:10.1140/epjc/s10052-008-0757-5, arXiv:0809.1567 [hep-ph].

[47] C. Anastasiou, L. J. Dixon, K. Melnikov and F. Petriello, *Phys. Rev.* **D69** (2004) 094008, doi:10.1103/PhysRevD.69.094008, arXiv:hep-ph/0312266.

[48] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, *JHEP* **04** (2004) 068, doi:10.1088/1126-6708/2004/04/068, arXiv:hep-ph/0303085.

[49] R. Hamberg, W. L. van Neerven and T. Matsuura, *Nucl. Phys.* **B359** (1991) 343, doi:10.1016/0550-3213(91)90064-5, [Erratum-*ibid.* **B644** (2002) 403].

[50] R. V. Harlander and W. B. Kilgore, *Phys. Rev. Lett.* **88** (2002) 201801, doi:10.1103/PhysRevLett.88.201801, arXiv:hep-ph/0201206.

[51] C. Anastasiou and K. Melnikov, *Nucl. Phys.* **B646** (2002) 220, doi:10.1016/S0550-3213(02)00837-4, arXiv:hep-ph/0207004.

[52] V. Ravindran, J. Smith and W. L. van Neerven, *Nucl. Phys.* **B665** (2003) 325, doi:10.1016/S0550-3213(03)00457-7, arXiv:hep-ph/0302135.

[53] R. V. Harlander and K. J. Ozeren, *JHEP* **11** (2009) 088, doi:10.1088/1126-6708/2009/11/088, arXiv:0909.3420 [hep-ph].

[54] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, JHEP **07** (2003) 001, doi:10.1088/1126-6708/2003/07/001, arXiv:hep-ph/0206293; *ALPGEN V2.14*, http://cern.ch/mlm/alpgen/.

[55] T. Gleisberg and S. Hoche, *JHEP* **12** (2008) 039, doi:10.1088/1126-6708/2008/12/039, arXiv:0808.3674 [hep-ph].

[56] E. Boos *et al.* [CompHEP Collaboration], *Nucl. Instrum. Meth.* **A534** (2004) 250, doi:10.1016/j.nima.2004.07.096, arXiv:hep-ph/0403113; *CompHEP homepage*, http://comphep.sinp.msu.ru/.

[57] A. Cafarella, C. G. Papadopoulos and M. Worek, *Comput. Phys. Commun.* **180** (2009) 1941, doi:10.1016/j.cpc.2009.04.023, arXiv:0710.2427 [hep-ph]; *HELAC-NLO and associated tools*, http://cern.ch/helac-phegas/.

[58] J. Alwall *et al.*, *JHEP* **09** (2007) 028, doi:10.1088/1126-6708/2007/09/028, arXiv:0706.2334 [hep-ph]; *The MadGraph5_aMC at NLO homepage*, http://madgraph.phys.ucl.ac.be/.

[59] E. W. N. Glover, Precision phenomenology and collider physics, talk given at Computer Algebra and Particle Physics Workshop, DESY Zeuthen, April 2005, http://www-zeuthen.desy.de/theory/capp2005/.

[60] F. A. Berends and W. T. Giele, *Nucl. Phys.* **B306** (1988) 759, doi:10.1016/0550-3213(88)90442-7.

[61] L. J. Dixon, arXiv:hep-ph/9601359.

[62] S. Catani and M. H. Seymour, *Nucl. Phys.* **B485** (1997) 291, doi:10.1016/S0550-3213(96)00589-5, [Erratum-*ibid.* **B510** (1998) 503], arXiv:hep-ph/9605323.

[63] S. Frixione, Z. Kunszt and A. Signer, *Nucl. Phys.* **B467** (1996) 399, doi:10.1016/0550-3213(96)00110-1, arXiv:hep-ph/9512328.

[64] D. A. Kosower, *Phys. Rev.* **D57** (1998) 5410, doi:10.1103/PhysRevD.57.5410, arXiv:hep-ph/9710213; J. M. Campbell, M. A. Cullen and E. W. N. Glover, *Eur. Phys. J.* **C9** (1999) 245, doi:10.1007/s100529900034, arXiv:hep-ph/9809429; D. A. Kosower, *Phys. Rev.* **D71** (2005) 045016, doi:10.1103/PhysRevD.71.045016, arXiv:hep-ph/0311272.

[65] Z. Nagy, *Phys. Rev.* **D68** (2003) 094002, doi:10.1103/PhysRevD.68.094002, arXiv:hep-ph/0307268; *NLOJet++*, https://www.desy.de/~znagy/Site/NLOJet++.html.

[66] J. M. Campbell and R. K. Ellis, *Phys. Rev.* **D62** (2000) 114012, doi:10.1103/PhysRevD.62.114012, arXiv:hep-ph/0006304; *MCFM - Monte Carlo for FeMtobarn processes*, http://mcfm.fnal.gov/.

[67] K. Arnold *et al.*, *Comput. Phys. Commun.* **180** (2009) 1661, doi:10.1016/j.cpc.2009.03.006, arXiv:0811.4559 [hep-ph]. *VBFNLO - A parton level Monte Carlo for processes with electroweak bosons*, https://www.itp.kit.edu/vbfnlo//wiki/doku.php.

[68] T. Binoth, J. P. Guillet, E. Pilon and M. Werlen, *Eur. Phys. J.* **C16** (2000) 311, doi:10.1007/s100520050024, arXiv:hep-ph/9911340; *The PHOX Family*, https://lapth.cnrs.fr/PHOX_FAMILY/main.html.

[69] A. Bredenstein, A. Denner, S. Dittmaier and S. Pozzorini, *Phys. Rev. Lett.* **103** (2009) 012002, doi:10.1103/PhysRevLett.103.012002, arXiv:0905.0110 [hep-ph].

[70] G. Bevilacqua, M. Czakon, C. G. Papadopoulos, R. Pittau and M. Worek, *JHEP* **09** (2009) 109, doi:10.1088/1126-6708/2009/09/109, arXiv:0907.4723 [hep-ph].

[71] G. Bevilacqua, M. Czakon, C. G. Papadopoulos and M. Worek, *Phys. Rev. Lett.* **104** (2010) 162002, doi:10.1103/PhysRevLett.104.162002, arXiv:1002.4009 [hep-ph].

[72] C. F. Berger *et al.*, *Phys. Rev.* **D80** (2009) 074036, doi:10.1103/PhysRevD.80.074036, arXiv:0907.1984 [hep-ph].

[73] R. Keith Ellis, K. Melnikov and G. Zanderighi, *Phys. Rev.* **D80** (2009) 094002, doi:10.1103/PhysRevD.80.094002, arXiv:0906.1445 [hep-ph].

[74] C. F. Berger *et al.*, *Phys. Rev.* **D82** (2010) 074002, doi:10.1103/PhysRevD.82.074002, arXiv:1004.1659 [hep-ph].

[75] C. F. Berger *et al.*, *Phys. Rev. Lett.* **106** (2011) 092001, doi:10.1103/PhysRevLett.106.092001, arXiv:1009.2338 [hep-ph].

[76] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover and G. Heinrich, *JHEP* **12** (2007) 094, doi:10.1088/1126-6708/2007/12/094, arXiv:0711.4711 [hep-ph].

[77] S. Weinzierl, *Phys. Rev. Lett.* **101** (2008) 162001, doi:10.1103/PhysRevLett.101.162001, arXiv:0807.3241 [hep-ph]; *JHEP* **06** (2009) 041, doi:10.1088/1126-6708/2009/06/041, arXiv:0904.1077 [hep-ph].

[78] K. Melnikov and F. Petriello, *Phys. Rev.* **D74** (2006) 114017, doi:10.1103/PhysRevD.74.114017, arXiv:hep-ph/0609070.

[79] S. Catani, L. Cieri, G. Ferrera, D. de Florian and M. Grazzini, *Phys. Rev. Lett.* **103** (2009) 082001, doi:10.1103/PhysRevLett.103.082001, arXiv:0903.2120 [hep-ph].

[80] C. Anastasiou, K. Melnikov and F. Petriello, *Nucl. Phys.* **B724** (2005) 197, doi:10.1016/j.nuclphysb.2005.06.03, arXiv:hep-ph/0501130.

[81] S. Catani and M. Grazzini, *Phys. Rev. Lett.* **98** (2007) 222002, doi:10.1103/PhysRevLett.98.222002, arXiv:hep-ph/0703012.

[82] T. Sjostrand, S. Mrenna and P. Skands, *Comput. Phys. Commun.* **178** (2008) 852, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820 [hep-ph]; *Present program version: PYTHIA 8.2*, http://home.thep.lu.se/~torbjorn/Pythia.html.

[83] T. Sjostrand, S. Mrenna and P. Skands, *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175; *PYTHIA 6.4*, https://pythiasix.hepforge.org.

[84] L. Lonnblad, *Comput. Phys. Commun.* **71** (1992) 15, doi:10.1016/0010-4655(92)90068-A.

[85] T. Gleisberg, S. Hoche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert and J. Winter, *JHEP* **02** (2009) 007, doi:10.1088/1126-6708/2009/02/007, arXiv:0811.4622 [hep-ph]; *Sherpa Homepage*, https://sherpa-team.gitlab.io/.

[86] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna and E. Norrbin, *Comput. Phys. Commun.* **135** (2001) 238, doi:10.1016/S0010-4655(00)00236-8, arXiv:hep-ph/0010017.

[87] G. Corcella *et al.*, *JHEP* **01** (2001) 010, doi:10.1088/1126-6708/2001/01/010, arXiv:hep-ph/0011363; *The Herwig Event Generator*, https://herwig.hepforge.org/.

[88] M. Bahr *et al.*, *Eur. Phys. J.* **C58** (2008) 639, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883 [hep-ph]; *The Herwig Event Generator*, http://projects.hepforge.org/herwig/.

[89] B. Andersson, G. Gustafson, G. Ingelman and T. Sjostrand, *Phys. Rep.* **97** (1983) 31, doi:10.1016/0370-1573(83)90080-7.

[90] T. Sjostrand, *Nucl. Phys.* **B248** (1984) 469, doi:10.1016/0550-3213(84)90607-2.

[91] B. R. Webber, *Nucl. Phys.* **B238** (1984) 492, doi:10.1016/0550-3213(84)90333-X.

[92] A. Buckley, H. Hoeth, H. Lacker, H. Schulz and J. E. von Seggern, *Eur. Phys. J.* **C65** (2010) 331, doi:10.1140/epjc/s10052-009-1196-7, arXiv:0907.2973 [hep-ph].

[93] T. Sjostrand, Monte Carlo Generators, 2006 European School of High-Energy Physics, Aronsborg, Sweden, Ed. R. Fleischer, CERN-2007-005 (CERN, Geneva, Switzerland, 2007), p. 51, doi:10.5170/CERN-2007-005.51, arXiv:hep-ph/0611247.

[94] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, JHEP **11** (2001) 063, doi:10.1088/1126-6708/2001/11/063, arXiv:hep-ph/0109231.

[95] J. Alwall *et al.*, *Eur. Phys. J.* **C53** (2008) 473, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569 [hep-ph].

[96] N. Lavesson and L. Lonnblad, *JHEP* **04** (2008) 085, doi:10.1088/1126-6708/2008/04/085, arXiv:0712.2966 [hep-ph].

[97] W. T. Giele, D. A. Kosower and P. Z. Skands, *Phys. Rev.* **D78** (2008) 014026, doi:10.1103/PhysRevD.78.014026, arXiv:0707.3652 [hep-ph].

[98] V. M. Abazov *et al.* [D0 Collaboration], *Phys. Lett.* **B678** (2009) 45, doi:10.1016/j.physletb.2009.05.058, arXiv:0903.1748 [hep-ex].

[99] S. Frixione and B. R. Webber, *JHEP* **06** (2002) 029, doi:10.1088/1126-6708/2002/06/029, arXiv:hep-ph/0204244.

[100] P. Nason, *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[101] P. Torrielli and S. Frixione, *JHEP* **04** (2010) 110, doi:10.1007/JHEP04(2010)110, arXiv:1002.4293 [hep-ph].

[102] O. Latunde-Dada, *JHEP* **05** (2009) 112, doi:10.1088/1126-6708/2009/05/112, arXiv:0903.4135 [hep-ph].

[103] S. Frixione, F. Stoeckli, P. Torrielli and B. R. Webber, *JHEP* **01** (2011) 053, doi:10.1007/JHEP01(2011)053, arXiv:1010.0568 [hep-ph].

[104] S. Alioli, P. Nason, C. Oleari and E. Re, *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581 [hep-ph].

[105] S. Frixione, P. Nason and B. R. Webber, *JHEP* **08** (2003) 007, doi:10.1088/1126-6708/2003/08/007, arXiv:hep-ph/0305252.

[106] G. P. Salam, *Eur. Phys. J.* **C67** (2010) 637, doi:10.1140/epjc/s10052-010-1314-6, arXiv:0906.1833 [hep-ph].

[107] S. D. Ellis, private communication to the OPAL Collaboration; D. E. Soper and H.-C. Yang, private communication to the OPAL Collaboration; L. A. del Pozo, PhD thesis, University of Cambridge, RALT–002, 1993; R. Akers *et al.* [OPAL Collaboration], *Z. Phys.* **C63**, (1994) 197, doi:10.1007/BF01411011.

[108] G. C. Blazey *et al.*, arXiv:hep-ex/0005012.

[109] G. P. Salam and G. Soyez, *JHEP* **05** (2007) 086, doi:10.1088/1126-6708/2007/05/086, arXiv:0704.0292 [hep-ph].

[110] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock and B. R. Webber, *Phys. Lett.* **B269**, (1991) 432, doi:10.1016/0370-2693(91)90196-W.

[111] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, *Nucl. Phys.* **B406** (1993) 187, doi:10.1016/0550-3213(93)90166-M.

[112] S. D. Ellis and D. E. Soper, *Phys. Rev.* **D48**, (1993) 3160, doi:10.1103/PhysRevD.48.3160, arXiv:hep-ph/9305266.

[113] M. Cacciari and G. P. Salam, *Phys. Lett.* **B641** (2006) 57, doi:10.1016/j.physletb.2006.08.037, arXiv:hep-ph/0512210.

[114] M. Cacciari, G. P. Salam and G. Soyez, *FastJet*, http://fastjet.fr/.

[115] M. Cacciari, G. P. Salam and G. Soyez, *JHEP* **04** (2008) 005, doi:10.1088/1126-6708/2008/04/005, arXiv:0802.1188 [hep-ph].

[116] M. Cacciari and G. P. Salam, *Phys. Lett.* **B659** (2008) 119, doi:10.1016/j.physletb.2007.09.077, arXiv:0707.1378 [hep-ph].

[117] V. M. Abazov *et al.* [D0 Collaboration], *Phys. Rev.* **D65**, 052008 (2002), doi:10.1103/PhysRevD.65.052008, arXiv:hep-ex/0108054.

[118] M. H. Seymour, *Z. Phys.* **C62** (1994) 127, doi:10.1007/BF01559532.

[119] J. M. Butterworth, B. E. Cox and J. R. Forshaw, *Phys. Rev.* **D65** (2002) 096014, doi:10.1103/PhysRevD.65.096014, arXiv:hep-ph/0201098.

[120] J. M. Butterworth *et al.*, arXiv:1003.1643 [hep-ph].

[121] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, *JHEP* **08** (1997) 001, doi:10.1088/1126-6708/1997/08/001, arXiv:hep-ph/9707323.

[122] M. Wobisch and T. Wengler, arXiv:hep-ph/9907280;
M. Wobisch, Diploma thesis, RWTH Aachen University, DESY-THESIS-2000-049, https://inspirehep.net/literature/538471.

[123] M. Cacciari, G. P. Salam and G. Soyez, *JHEP* **0804** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189 [hep-ph].

[124] M. Cacciari, J. Rojo, G. P. Salam and G. Soyez, *JHEP* **0812** (2008) 032, doi:10.1088/1126-6708/2008/12/032, arXiv:0810.1304 [hep-ph].

[125] M. Cacciari, J. Rojo, G. P. Salam and G. Soyez, *Testing jet definitions: qq and gg cases*, http://quality.fastjet.fr.

[126] J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, *Phys. Rev. Lett.* **100** (2008) 242001, doi:10.1103/PhysRevLett.100.242001, arXiv:0802.2470 [hep-ph].

[127] D. Krohn, J. Thaler and L. T. Wang, *JHEP* **02** (2010), 084, doi:10.1007/JHEP02(2010)084, arXiv:0912.1342 [hep-ph].

[128] C. Buttar *et al.*, arXiv:0803.0678 [hep-ph].

[129] D. Krohn, J. Thaler and L. T. Wang, *JHEP* **06** (2009), 059, doi:10.1088/1126-6708/2009/06/059, arXiv:0903.0392 [hep-ph].

[130] M. Dasgupta, L. Magnea and G. P. Salam, JHEP **02** (2008) 055, doi:10.1088/1126-6708/2008/02/055 arXiv:0712.3014 [hep-ph].

[131] M. Rubin, *JHEP* **05** (2010), 005, doi:10.1007/JHEP05(2010)005 arXiv:1002.4557 [hep-ph].

[132] 2008 CTEQ - MCnet Summer School on QCD Phenomenology and Monte Carlo Event Generators, Debrecen, Hungary, 2008, https://www.physics.smu.edu/scalise/cteq/schools/summer08/.

[133] CTEQ Summer School on QCD Analysis and Phenomenology, Madison, WI, USA, 2009, https://www.physics.smu.edu/scalise/cteq/schools/summer09/.