Proceedings of the 2018 European School of High-Energy Physics, Maratea, Italy, 20 June–3 July 2018
Edited by M. Mulders and C. Duhr, CERN-2019-006 (CERN, Geneva, 2019)

Introduction to perturbative QCD

F. Tramontano
Università di Napoli Federico II e INFN sezione di Napoli, Napoli, Italia

Abstract
Following a phenomenological guiding principle, in these notes I present a selection of old and recent developments in perturbative QCD and collider physics.

Keywords
Quantum Chromodynamics; lectures; gauge theory; hadrons; parton shower.

1 Basics and e^+e^- annihilation into hadrons

Protons, neutrons and hadrons mainly interact via the so called strong force. The strong force has a higher level of symmetry than weak and electromagnetic interactions, since it conserves parity and isospin (almost). Unlikely the case of QED, there is no evidence of a low energy small parameter which controls the interaction strength. On the contrary, the strong interaction is characterized by a dimensional scale of about $200 \sim 300\,\mathrm{MeV}$, corresponding to the typical lifetime of its excitations that is about $10^{-24}\,\mathrm{s}$. As can be seen in Fig.(1), taken from the Review of Particle Physics [1], the typical hadron-hadron cross section is of the order of $10\,\mathrm{mb}$, that again corresponds to about $(300\,\mathrm{MeV})^2$, so it's very much harder

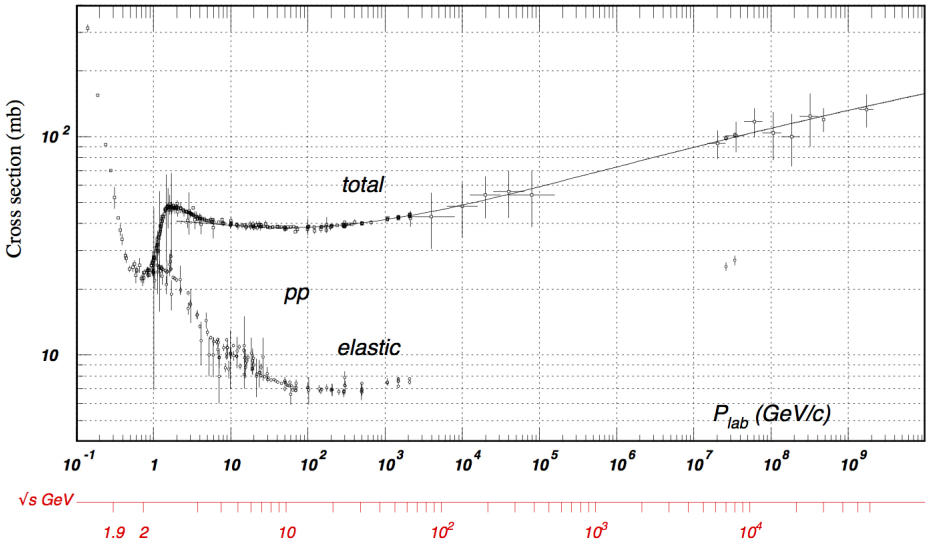


Fig. 1: Measurements of the total proton-proton cross sections as function of the beam energy.

then EW interactions and furthermore we don't really see the vertices.

http://doi.org/10.23730/CYRSP-2019-006.43

1.1 Motivations for QCD

First of all, the whole hadron spectrum can be classified by assuming that the hadrons are made up of quarks carring different flavours (see Fig.(2) [1]), that each quark q with flavour f comes in three

quark flavour	spin	charge (e)
d=down	1/2	-1/3
u=up	1/2	2/3
s=strange	1/2	-1/3
c=charm	1/2	2/3
b=beauty	1/2	-1/3
t=top	1/2	2/3

Fig. 2: Quark flavours and their properties.

different colours (q_f^i with index $i = 1, 2, 3$) and that observable hadrons are colour singlets under the $SU(3)_C$ colour group. Singlets under the $SU(3)_C$ are easily built for a quark-antiquark (meson) or a three quark system (baryon). In Fig.(3) [1], for example, there are the lowest mass multiplets classified as representation of favour $SU(4)_F$. There is a nice symmetry in each group and mass differences are well

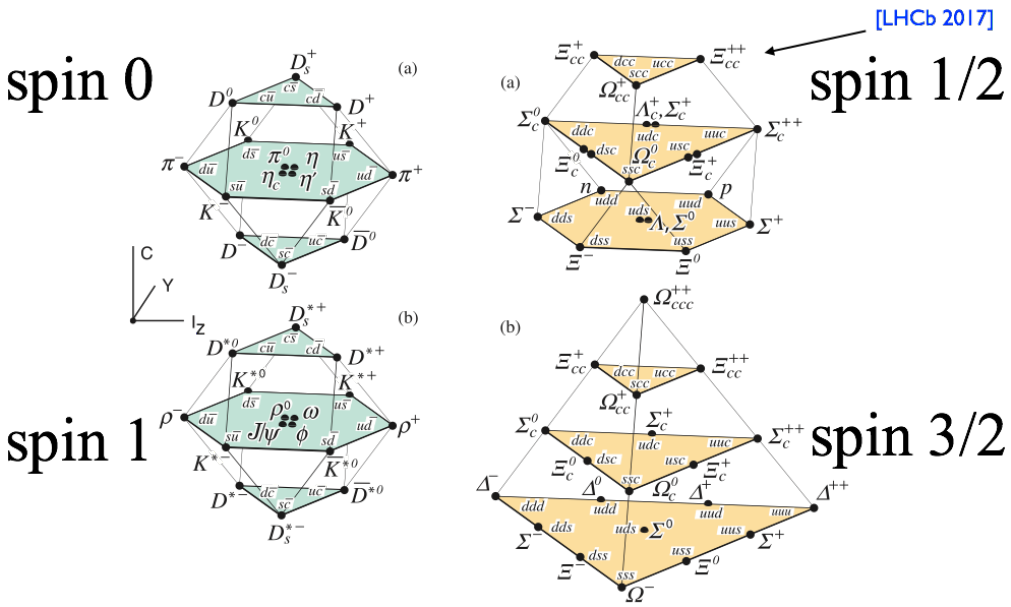


Fig. 3: Classification of the light mesons and baryons spectra according to the multiplets of the $SU(4)_F$ flavour group.

explained by quark masses. The relatively low number of multiplets and their composition is difficult to explain without colour constraint. For example, consider the Δ^{++} state, that is composed of three up quarks with their spin aligned. For the Pauli principle this state would not be possible without assuming total antisymmetry of its colour wave function. One could in principle assume an antisymmetric space wave function, but this would naturally require to this state to have a mass larger than the others in the same multiplet that have a different flavour content so that for them a symmetric wave function is natural; on the contrary they are all close by. It's remarkable that the Ξ_{cc}^{++}, the last missing state in the proton

multiplet to be discovered, has been found only in 2017 by the LHCb experiment [2]. However, the simple vision of just $q\bar{q}$ and qqq states does not represent the whole story. There is now evidence of states that do not fit the spectrum dictated by these two combinations only, both for the meson and the baryon sector. The new states found involve the presence of both heavy (c or b) or light (u, d or s) quarks, so that there are light-light, heavy-light and heavy-heavy new resonances. For example, in the light-light sector, there is observation of $\pi\pi$ resonances with isospin 2 while quarks have only isospin $1/2$, and for the hidden charm and beauty sectors (resonances containing a $c\bar{c}$ or a $b\bar{b}$ pair), there is evidence of many states that for their masses, decay modes and decay widths do not fit at all the spectrum of heavy $q\bar{q}$ resonances (quarkonium). All such states have to be necessarily multi-quark (more than three) systems! However, the only necessary condition to build a colour singlet is that:

$$n_q - n_{\bar{q}} = \text{ multiple of } 3 \tag{1}$$

which also implies the absence of hadrons with fractional charge although quarks have fractional charges. At the same time Eq.(1) predicts the existence of multi-quark colourless states. So far, indeed, QCD has more then one solution to the problem! See the Fig.(4). The new states could resemble nuclei, being

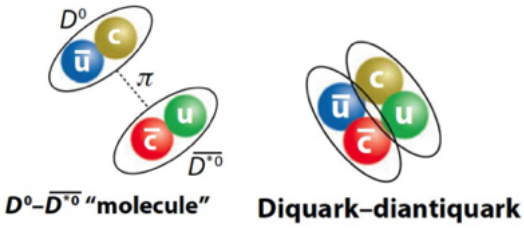


Fig. 4: Two QCD allowed solutions for a multi-quark state: a "molecule" made of regular hadrons binded together by the nuclear force (left) and new exotic tetra-quark (right).

molecules of color singlet states kept together by kind of nuclear forces: that would explain the fact that these states have masses typically close to the respective thresholds. Or they could also be made of colored parts confined by long-range color forces, similarly to what happens in normal mesons, making tetraquarks (or pentaquarks, hexaquarks...) as indicated by the spectrum of the $X(3872)$ (a 1^{++} meson decaying into $J/\psi\,\pi^+\,\pi^-$) measured by CMS, that is much harder than the (rescaled) deuteron and tritium spectrum measured in Pb-Pb collisions, see Fig.(5) [3]. We refrain here to discuss this aspects further, nevertheless we stress that such studies might indeed provide new clues to the understanding of QCD in the fully non-perturbative regime.

The second pillar in favour of QCD is the phenomenon of scaling in lepton-hadron Deep Inelastic Scattering (DIS). By this, we mean the fact that, for inclusive inelastic large angle scattering, the differential cross section, as expressed in terms of dimensionless variables, "scales" with energy:

$$\frac{d\sigma}{dxdy} \sim \frac{1}{s} \qquad (s = E_{CM}^2) \tag{2}$$

As we will discuss further, this is a prediction of the parton model of QCD. So, the strong interaction at high energy resembles a weakly interacting theory with dimensionless coupling!

Last but not the least, there has been the theoretical discovery of asymptotically free theories. Non abelian gauge theories are weakly coupled at high energies (short distances) while they can become strongly approaching a characteristic low energy scale. In the latter highly non perturbative regime, the strong interacting fields may lead to a spectrum of asymptotic binding states, "confining" the high energy degrees of freedom. So far colour has been used to justify the spectrum that is a static property of the theory, nevertheless assuming that the outcome of every strong interaction are only colourless particles

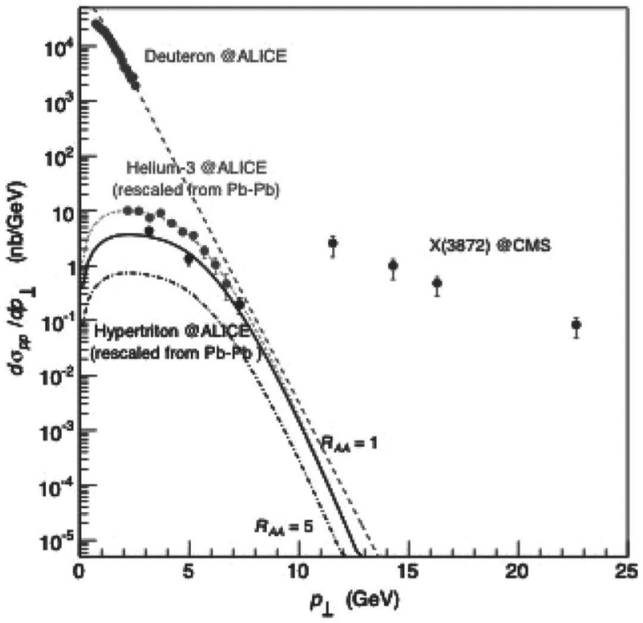


Fig. 5: Comparison of the measureed transverse momentum spectrum of $X(3872)$ in p-p collision (CMS) with the one of deuteron and tritium in Pb-Pb collisions (ALICE). The huge discrepancy disfavors the interpretation of the new state $X(3872)$ as a molecule.

we are implicitly also giving to colour a dynamical role. For all these reasons, non abelian gauge theories ($SU(3)_C$ in our case) are good candidates to be the theory of strong interactions. One has to assume that the strong interaction group is completely independent from the weak interaction group (commute) to guarantee that parity violating terms remain of order α_e/M_W^2, but still one has same type of theory for weak, strong and electromagnetic interactions! The basic aspects of QCD are discussed in excellent books like for eample [4] and [5].

1.2 QCD Lagrangian and Feynman rules

The QCD Lagrangian is given by:

$$\mathcal{L} = -\frac{1}{4}F^A_{\mu\nu}F_{\mu\nu A} + \sum_{f=1}^{N_f}\bar{q}^i_f(i\gamma^\mu\partial_\mu\delta_{ij} - g_s\gamma^\mu t^A_{ij}A^A_\mu - m_f\delta_{ij})q^j_f. \tag{3}$$

For $N_c = 3$ colours we have $i, j = 1, 2, 3$ and $A, B, C = 1, 2, ..., 8$. The field strength is

$$F^A_{\mu\nu} = \partial_\mu A^A_\nu - \partial_\nu A^A_\mu - g_s f_{ABC}A^B_\mu A^C_\mu. \tag{4}$$

The eight SU(3) generators fulfill the commutation rules:

$$[t^A, t^B] = if_{ABC}t^C \tag{5}$$

where f_{ABC} are numbers specific of the SU(3) group named the structure constants of the group. The terms in the gluon field strength generate 3-gluon and 4-gluon couplings. Then, unlike the neutral single photon in the case of QED, the eight gluons carry colour charge. Furthermore, gluon radiation from a quark or a gluon changes its colour charge generating a colour flow. The explicit colour matrices are not

important for most practical purposes. It is customary to normalise them according to

$$Tr(t^A t^B) = T_R \delta^{AB} \qquad T_R = \frac{1}{2} \tag{6}$$

and use $t^A = \frac{1}{2}\lambda^A$, where λ^A are the Gell-Mann matrices:

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda^3 = ... \tag{7}$$

From this Lagrangian one finds the Feynman rules given in Fig.(6). Before using them for the computa-

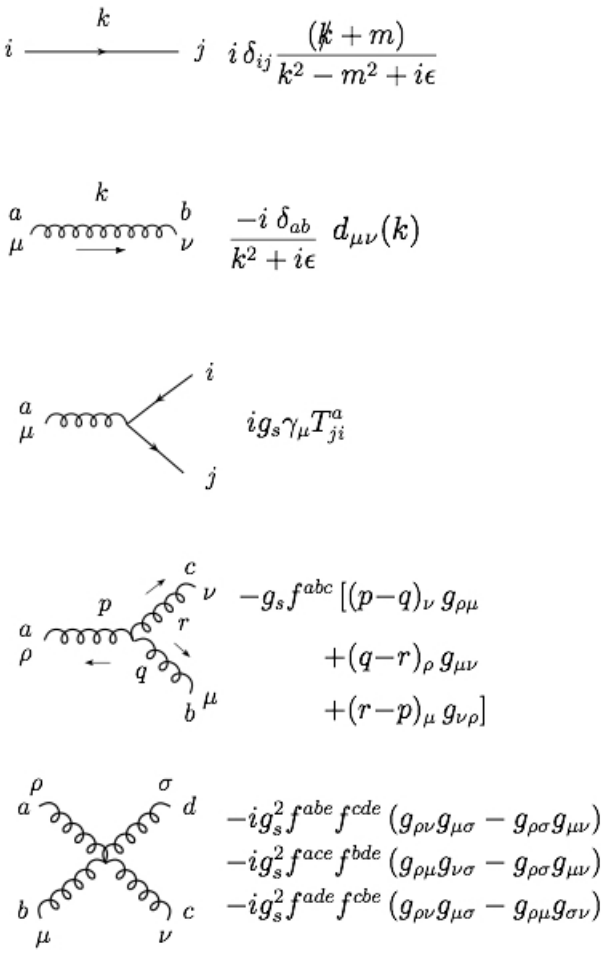


Fig. 6: List of the Feynman rules corresponding to the QCD Lagrangian in Eq. 3.

tion of amplitudes in perturbation theory, one has to specify the gluon polarization tensor $d_{\mu\nu}(k)$ which appears in the Feynman rule for the gluon propagator. Several choices are possible, for example:

$$d_{\mu\nu}(k) = \sum_\lambda \varepsilon^\mu_\lambda(k)\varepsilon^{\nu *}_\lambda(k) = \begin{cases} -g^{\mu\nu} + (1-\rho)\frac{k^\mu k^\nu}{k^2+i\epsilon} & \text{covariant} \\ \\ -g^{\mu\nu} + \frac{n_\mu k_\nu + k_\mu n_\nu}{n\cdot k} - \frac{(n^2-\rho k^2)k^\mu k^\nu}{(n\cdot k)^2} & \text{axial} \end{cases} \tag{8}$$

With ρ and n a free parameter and fixed four vector respectively, and the property that physical results have to be independent from them. The easiest and most natural choice seems to be the covariant gauge

with $\rho = 1$. However, it propagates also unphysical (non transverse) degrees of freedom, as a consequence of the non-abelian nature of the theory. This spurious degrees of freedom have to be ruled out in the calculation of physical quantities. The simplest way to implement this cancellation is via the on purpose introduction of other non physical degrees of freedom named ghosts. So, if one adopts the covariant gauge, the Feynman rules have to be extended including the complementary ones shown in Fig.(7). To meet the purpose, ghosts are coloured scalars with a fermionic property: a (-1) factor for each

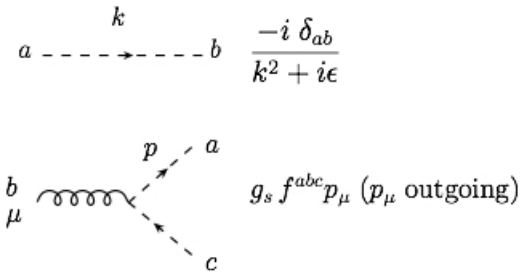


Fig. 7: Feynman rules involving ghost particles in covariant gauge.

ghost loop has to be included as, indeed, for the fermion loop. The axial gauges, instead, propagates only the physical degrees of freedom, and so ghosts are not required, to the price of a more complicated expression in the propagator. Let us focus on a specific example. Suppose you want to compute the scattering amplitude for the process $q\bar{q} \to gg$ that is one of the subprocesses at work to create two jets at an hadron collider. The corresponding Feynman diagrams, which involve only the physical degrees of freedom, are the ones in Fig.(8). The cross section is proportional to the squared amplitude, and one can

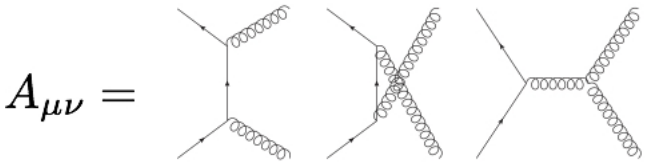


Fig. 8: Lowest order Feynman diagrams contribution to the scattering amplitude for $q\bar{q} \to gg$.

compute it according to the formula:

$$\sigma = A_{\mu\nu} A^*_{\mu'\nu'} d^{\mu\mu'} d^{\nu\nu'} \tag{9}$$

generically written in terms of the polarization tensors for the two gluons. If you decide to use the gluon polarization tensor as given in the axial guage there is nothing more to add. The point is that if one uses naively the simple covariant expression, i.e. $d^{\alpha\beta} = -g^{\alpha\beta}$, to compute the squared amplitude in Eq.(9), will end up with a wrong result. The correct one is recovered after adding also the ghost contribution shown in Fig.(9). What is done in several automated programs designed to calculate scattering

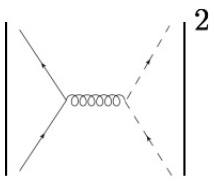


Fig. 9: Ghost contribution to the matrix element for the $q\bar{q} \to gg$ scattering process in a covariant gauge

amplitudes is to compute the helicity amplitudes, that is the collection of amplitudes each with a specific

choice for the helicities of the external particles. In this case one considers ab initio only physically meaningful helicity states. Then the square is obtained using the formula:

$$\sigma = \sum_{\substack{helicity \\ configurations}} |A_{h.c.}|^2 \tag{10}$$

Numerically, this procedure is very efficient. Beyond Dirac algebra we also need to compute the colour factors associated to the colour structure of the QCD vertices. The most relevant ones are related to the computations showed in Fig.(10). As we see from the second and third line in Fig.(10), the gluon

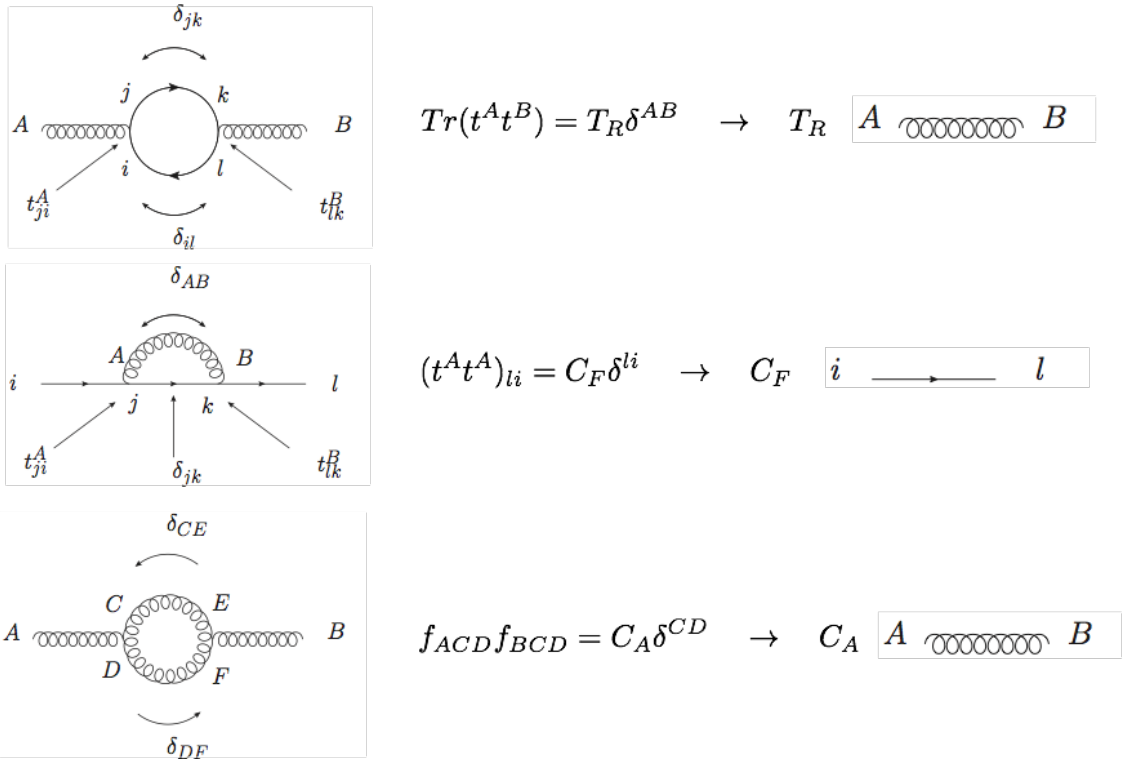


Fig. 10: Examples of fundamental colour structures and associated colour factors.

radiative correction to the fermion propagator receives a factor $C_F = 4/3$ while the correction to the gluon propagator receives a factor $C_A = 3$. From this difference, we argue that gluons radiate more then fermions, their coupling is effectively almost twice as much as the one of the fermions. Another colour identity that is very useful in practice is the Fierz transformation:

$$t^A_{ij} t^A_{kl} = \frac{1}{2}\delta_{kj}\delta_{il} - \frac{1}{2\,N_c}\delta_{ij}\delta_{kl} \tag{11}$$

which allows one to deal with colour flowing from one line to another as in Fig.(11).

In principle, starting from the Feynman diagrams, one has that a helicity amplitude can be written as a combination of the colour tensor structures (C_i), that are just strings of colour matrices, and fully contracted Lorentz structures (complex numbers K_{ij}) for each colour structure (i) and external helicity configurations (j)

$$A_j = \sum_i C_i \cdot K_{i,j}. \tag{12}$$

The same numerical construction can be generalized also for diagrams containing one closed loop. Indeed, we can factorize one loop diagrams into kinematical factors and integrals over the unconstrained

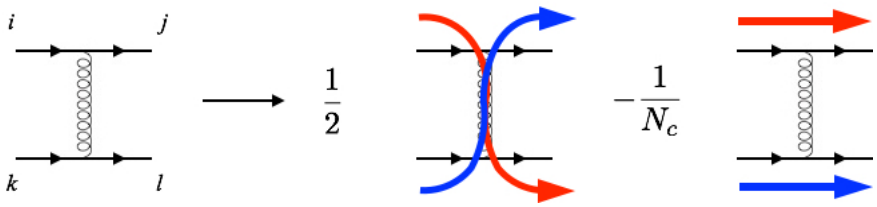


Fig. 11: Diagrammatic colour flow in a quark-quark scattering.

loop momentum. Note that possible one loop integrals (there is finite number of them) are all known analytically so that, beyond Dirac algebra, what is needed to build the "interference" among tree level and one loop diagrams, that is part of the Next-to-Leading Order (NLO) correction, is just the colour matrix built by interfering (contracting) all the colour structures present in both. Full automation of tree level and one loop amplitude computation has been reached in recent years. We talk of the "NLO revolution" to express the fact that, based on a large number of optimizations in the computation of the kinematical factors ($K_{i,j}$) and the availability of faster computers and larger memories with respect to the past, nowadays it is possible to generate full codes for very complicated processes with up to hundreds of thousands of Feynman diagrams just pushing a button and waiting (only once) an amount of time that is much lower then the running time needed to make simulations with that code with sufficient numerical precision. Although, as we will see in another section, the NLO revolution has a prequel.

1.3 QCD Phenomenology in e^+e^- collisions

We now start the phenomenological exploration of QCD. It will be clearer in the following that we have to distinguish among different regimes: at low energy QCD studies address hadron spectrum as a consequence of the symmetries, the properties of finite temperature hadronic matter and lattice calculations. We will instead focus on the high energy phenomenology: in particular on perturbation theory with quarks and gluons, although only hadrons appear in the initial and final states of the scattering processes. The assumption here is that the processes of quark extraction from an initial hadron and that of hadron formation do not spoil the predictions that one can perform working with quarks and gluons. This has implications on the definition of observables that can be predicted in perturbation theory. Other fundamental applications of QCD are connected with the effective field theories. They cover several aspects and topics which play a very important role both to validate QCD and to establish QCD effects on processes that are beyond the Standard Model. This is a field that is relevant at all energies, nevertheless we will not treat it here as it would require more dedicated lectures.
Let's start from the easiest situation that is represented by the hadron production in electron-positron collisions, namely the process $e^+e^- \to \gamma^* \to$ hadrons. For energies below the Z pole, the prediction of QCD at the lowest order in the strong coupling constant, i.e. $\alpha_s^0 = (g_s^2/(4\,\pi))^0$, for the ratio among the cross sections for hadrons and muon pair creation is:

$$R = \frac{\sigma(\gamma^* \to hadrons)}{\sigma(\gamma^* \to \mu^+\mu^-)} = N_c \sum_i q_i^2 = 3 \cdot \left(\frac{4}{9} + \frac{1}{9} + \frac{1}{9} + \frac{4}{9} + \frac{1}{9} + ...\right) \tag{13}$$

that is obtained squaring the diagram in Fig.(12) and the rational numbers in parenthesis are the squared charges of u, d, s, c, b ... quarks (q_i). At the next order in QCD, i.e. at order α_s, one has to compute all the diagrams shown in Fig.(13), corresponding to the emission of a virutal or real gluon off the external quarks. They are indeed the same as if the corrections would have been induced by a QED photon, but for a colour factor of $C_F = 4/3$ (in QED the correction is $3\alpha_e/4$). At order α_s^2, instead, we encounter a

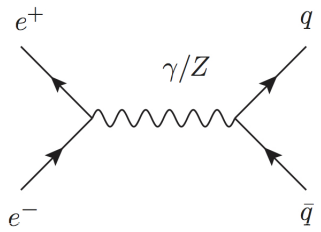


Fig. 12: Lowest order diagram to the process $e^+e^- \to$ hadrons. An electron-positron pair annihilates into a photon or a Z boson which converts into quark anti-quark pair.

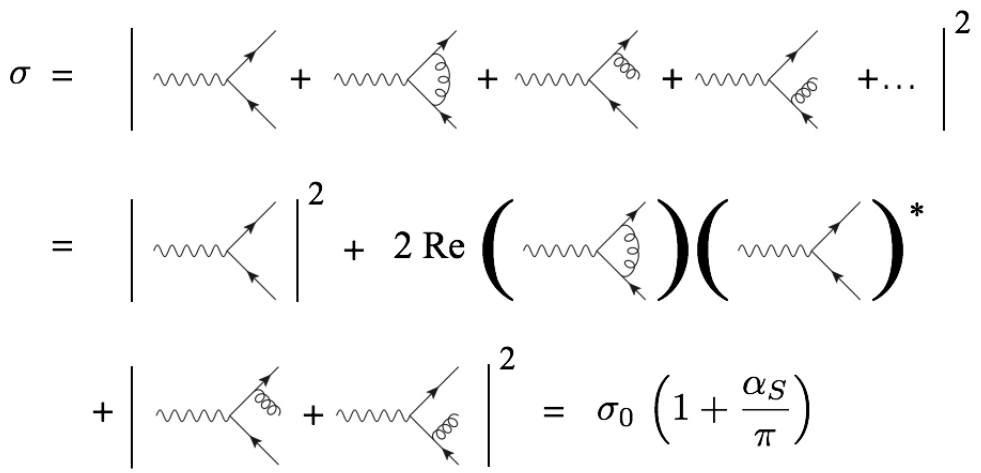


Fig. 13: At Next-to-Leading Order in the QCD interaction, the scattering amplitude receives contributions from both virtual and real emission of an extra gluon off the external quark lines. After squaring the amplitude and integrating it, the correction amount to the leading order cross section σ_0 times $\frac{\alpha_s}{\pi}$.

problem. Once we add all the diagrams we get:

$$\sigma = \sigma_0 \left\{ 1 + \frac{\alpha_s}{\pi} + \left[c + \pi b_0 \log \frac{M^2}{Q^2} \right] \frac{\alpha_s^2}{\pi^2} \right\} \tag{14}$$

with

$$c = 1.986 - 0.115\, n_f \qquad b_0 = \frac{33 - 2\, n_f}{12\pi} \tag{15}$$

In this formula for the total cross section we have indeed a logarithmic ultraviolet divergence that we have regulated by introducing the ultraviolet cutoff energy scale M, while Q is the invariant mass of the system, i.e. the "hard" scale of the process, and n_f is the total number of quark flavours that can enter in the loops. More in general, one can show that for any physical quantity G computed as a power series in the coupling constant:

$$G = G_0 \alpha_s^n + (...)\alpha_s^{n+1} + ... \tag{16}$$

the expansion has the form:

$$G = G_0 \alpha_s^n + \left(G_1 + n\, G_0\, b_0\, \log \frac{M^2}{Q^2} \right) \alpha_s^{n+1} + ... \tag{17}$$

Indeed, for the e^+e^- cross section into hadrons up to second order Eq.(14) gives:

$$\frac{\sigma}{\sigma_0} - 1 = \frac{1}{\pi}\alpha_s + \left(\frac{c}{\pi^2} + \frac{1}{\pi} b_0 \log \left(\frac{M^2}{Q^2} \right) \right) \alpha_s^2 \tag{18}$$

where we can easily make the identifications: $n = 1$, $G_0 = 1/\pi$ and $G_1 = c/\pi^2$. Now we define:

$$\widetilde{\alpha}_s(\mu) = \alpha_s + b_0 \log\left(\frac{M^2}{\mu^2}\right)\alpha_s^2 \tag{19}$$

introducing a new fictitious scale μ. So that in terms of $\widetilde{\alpha}_s(\mu)$, neglecting terms of order α_S^{n+2} and higher, we find that any physical quantity gets the form:

$$G = G_0\widetilde{\alpha}_s^n(\mu) + \left(G_1 + nb_0 \log\left(\frac{\mu^2}{Q^2}\right)\right)\widetilde{\alpha}_s^{n+1}(\mu) + ... \tag{20}$$

where all the dependence on the cutoff regulator M has been adsorbed, and it is now contained, in the redefinition of the coupling constant. This is what is called renormalization.
More formally, we have that for any physical quantity $G(\alpha_s, M, ...)$ depending on the strong coupling, an ultraviolet cutoff, and physical variables like masses and momenta, we can always define a charge containing the whole dependence on the UV cutoff M:

$$\widetilde{\alpha}_s(\mu, M, \alpha_s) = \alpha_s + c_1(\mu, M)\alpha_s^2 + ... \tag{21}$$

in such a way that the physical quantity has a finite expression in terms of $\widetilde{\alpha}$, μ and the physical variables:

$$G\left(\alpha_s, M, ...\right) = G'\left(\widetilde{\alpha}_s(\mu, M, \alpha_s), \mu, ...\right). \tag{22}$$

Even if the definition of the coupling contains divergences, they are of ultraviolet origin and we might think that other dynamical effects may solve the problem at high energies while for the moment, at the energies we are probing the theory, we can assume the coupling to be finite, measure it in one process and use it to make predictions for another process! Before doing this, let us look more carefully to Eq.(22). Taking derivative on both sides wrt $\log(\mu^2)$ we have (note that G in the left hand side of Eq.(22) has no μ dependence):

$$0 = \frac{\partial G'\left(\widetilde{\alpha}_s, \mu, ...\right)}{\partial \widetilde{\alpha}_s}\frac{\partial \widetilde{\alpha}_s}{\partial \log(\mu^2)} + \frac{\partial G'\left(\widetilde{\alpha}_s, \mu, ...\right)}{\partial \log(\mu^2)} \tag{23}$$

or, equivalently:

$$\frac{\partial \widetilde{\alpha}_s}{\partial \log(\mu^2)} = -\frac{\frac{\partial G'(\widetilde{\alpha}_s, \mu, ...)}{\partial \log(\mu^2)}}{\frac{\partial G'(\widetilde{\alpha}_s, \mu, ...)}{\partial \widetilde{\alpha}_s}} = \beta\left(\widetilde{\alpha}_s, \mu\right) \tag{24}$$

But β is dimensionless, so if it does not depend explicitly on M it cannot depend explicitly on μ, so that:

$$\beta \equiv \beta\left(\widetilde{\alpha}_s\right) = \frac{\partial \widetilde{\alpha}_s}{\partial \log(\mu^2)} \tag{25}$$

Then, from e^+e^- cross section into hadrons we have:

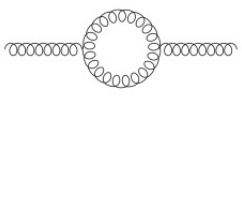
$$\widetilde{\alpha}_s(\mu) = \alpha_s + b_0 \log\left(\frac{M^2}{\mu^2}\right)\alpha_s^2 \quad \Rightarrow \quad \frac{\partial \widetilde{\alpha}_s}{\partial \log \mu^2} = -b_0\widetilde{\alpha}_s^2 + \mathcal{O}\left(\widetilde{\alpha}_s^3\right) \tag{26}$$

In summary, any physical quantity can be given as an expansion in $\widetilde{\alpha}_s$:

$$G'\left(\widetilde{\alpha}_s, \mu, ...\right) = \sum_i G_i(\mu, ...)\widetilde{\alpha}_s^i \tag{27}$$

and if we change μ and $\widetilde{\alpha}_s$ in such a way that:

$$\delta\widetilde{\alpha}_s = \beta(\widetilde{\alpha}_s)\delta \log(\mu^2) \tag{28}$$



$$\text{QCD:}\quad b_0 = \frac{33 - 2n_f}{12\pi} > 0$$

$$\text{QED:}\quad b_0 = -\frac{4n_f}{12\pi} < 0$$

Fig. 14: One loop contributions to the beta function in QCD and QED for the running of the coupling constant. While the fermionic loops gives raise to the same contribution in both cases (apart for the constant factor $C_F = 4/3$), the gluon bubble insertion in the QCD case turns the sign of the b_0 coefficient from negative to positive.

the prediction of physical quantities does not change. The last property goes under the name of Renormalization Group Invariance. After renormalization, the e^+e^- cross section into hadrons has become (dropping the tilde from now on):

$$\sigma = \sigma_0 \left\{ 1 + \frac{\alpha_s(\mu)}{\pi} + \left[c + \pi b_0 \log\left(\frac{\mu^2}{Q^2}\right)\right] \frac{\alpha_s^2(\mu)}{\pi^2} \right\} \tag{29}$$

It will be better to choose $\mu \sim Q$ otherwise the second order may become larger then the first (and subsequent orders even larger!). Remarkably, choosing $\mu = Q$ (or taking the two in a fixed ratio), the energy dependence of the cross section is entirely described by the scale dependence (running) of α_s and so it is of fundamental importance to solve the equation on the right hand side of Eq.(26). This equation can be recast in the following form:

$$\frac{\partial}{\partial \log \mu^2} \frac{1}{\alpha_s} = b_0 \tag{30}$$

with the obvious solution:

$$\frac{1}{\alpha_s} = b_0 \log \mu^2 + constant \tag{31}$$

and the integration constant can be used to get the expression:

$$\alpha_s(\mu) = \frac{1}{b_0 \log \frac{\mu^2}{\Lambda^2}} \tag{32}$$

The parameter Λ, or Λ_{QCD}, has nothing to do with a scale value where the coupling diverges. It is just a smart way to rewrite the integration constant. Furthermore, for scales where the coupling becomes large, one must include higher order corrections to the β function in Eq.(26) leading to a different structure of the solution. The logarithmic ultraviolet divergences come mainly from the bubble graphs in Fig.(14) and the main difference with the QED case is due to the gluon bubble that brings an opposite sign with respect to the contribution of the fermionic bubbles. In QCD, Λ has to be in the range of the typical hadronic scale we have mentioned above, the scale at which the hadronic systems become strongly coupled. Assuming Λ in the range $100 \div 500\,\mathrm{MeV}$, Eq.(32) gives $\alpha_s(M_Z) = 0.1 \div 0.13\,(\pm 13\%)$ and $\alpha_s(10^7\mathrm{GeV}) = 0.040 \div 0.044\,(\pm 5\%)$, i.e. going up with energy, the logarithmic dependence nicely narrows the relative uncertainty. The current status of the computation of e^+e^- annihilation into hadron is given by the formula:

$$\frac{\sigma}{\sigma_0} = 1 + \frac{\alpha_s}{\pi}\left(1 + 0.448\alpha_s - 1.30\alpha_s^2 - 2.59\alpha_s^3\right) + ... \tag{33}$$

where the last term has been computed in [6]. The numerical coefficients are obtained for the case $n_f = 5$ (the general expression is very complicated) and $\alpha_s \equiv \alpha_s^{\overline{MS}}(Q)$, with $\overline{MS}$ reflecting the details of the renormalization prescription used to define the coupling constant. Corrections are well behaved and amount to about 5% at the second order, 2% at the third order and $4‰$ at the fourth order. In principle

one could think to use this formula to determine α_s with an accuracy better then 1%, although the e^+e^- cross section is not an observable very sensitive to QCD effects since the perturbative expansion only starts at order α_s^0. As for the β function, the situation is as follows:

$$\frac{d\alpha_s}{d\log\mu^2} = -b_0\alpha_s^2 - b_1\alpha_s^3 - b_2\alpha_s^4 - b_3\alpha_s^5 - b_4\alpha_s^6 \tag{34}$$

where the last coefficient has been computed by three groups [7–9]. The first coefficients are:

$$b_0 = \frac{33-2n_f}{12\pi}, \quad b_1 = \frac{153-19n_f}{24\pi^2}, \quad b_2 = \frac{2857-\frac{5033n_f}{18}+\frac{325n_f^2}{54}}{(4\pi)^4} \tag{35}$$

In an actual computation at a given order in perturbation theory, the usual prescription is to use the running of the coupling constant at one higher order. Therefore, since there are not so many very higher order computations, what is used in general is the second order running:

$$\alpha_s(\mu) = \frac{1}{b_0\log\frac{\mu^2}{\Lambda^2}}\left(1-\frac{b_1\log\log\frac{\mu^2}{\Lambda^2}}{b_0^2\log\frac{\mu^2}{\Lambda^2}}\right) \tag{36}$$

Measurements of α_s have been done combining large data sets as the one reported in Fig.(15). The curvi-

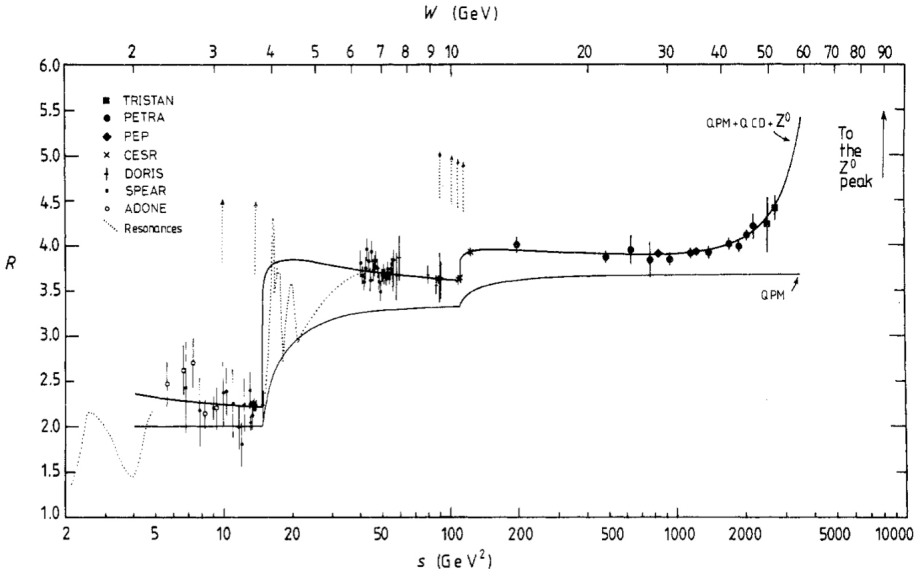


Fig. 15: A compilation of R ratio measurements at different energies used for the extraction of α_s.

linear behaviour in the prediction is given by a threshold velocity term to take into account that quarks are not massless. NLO corrections are clearly demanded for the fit. Combining results of experiments from 20 up to 65GeV, the value $\alpha_s(35\text{GeV}) = 0.146 \pm 0.030$ is found, that after evolution implies: $\alpha_s(M_Z) = 0.124 \pm 0.021$. Other ways to determine α_s is through the measurement of the total hadronic decay width at the Z pole at LEP, which gives $\alpha_s(M_Z) = 0.122 \pm 0.009$, or comparing hadronic and leptonic tau decays. In this case one has that QCD corrections displaces the value of the ratio from 3, that is a factor entirely due to colour. It is found $\alpha_s(M_\tau) = 0.36 \pm 0.05$ that after evolution implies: $\alpha_s(M_Z) = 0.122 \pm 0.005$.

1.4 How to compute more features of the final state

The nice agreement of the coupling constants extracted from total cross sections and decay rates mentioned at the end of the previous section is certainly an indication of the goodness of the theory, but how

to go further and prove the validity of the theory also at differential level? Indeed, to compute e^+e^- annihilation into hadrons at the first order in the strong coupling, we have combined diagrams for producing two quarks and two quarks and a gluon, all particles that btw have never been observed. Furthermore, the final state partons bring colour quantum number that has never been observed neither. It is clear that we have to find the conditions that allow for the description of the final state in terms of an evolution from the elementary processes (among quark and gluons) to the hadrons that are observed. In doing that we will also gain the possibility to make differential predictions with QCD. We will again use the radiative corrections for e^+e^- to hadrons to give an idea of such a procedure. As we have seen in Fig.(13), the first perturbative correction to e^+e^- annihilation into hadrons (NLO QCD) is computed adding real and virtual emission diagrams. The technical difficulty is that the contributions live in different phase spaces. We start with the computation of the diagrams for the real emission in Fig.(16).

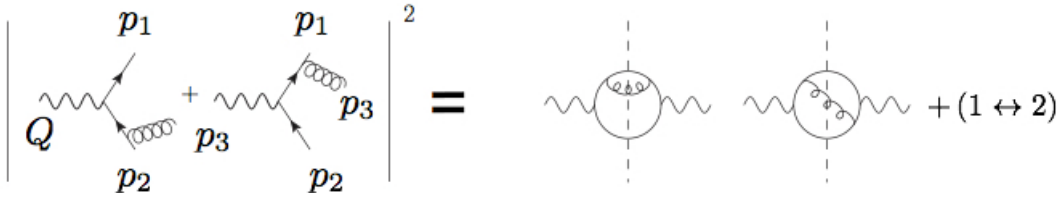


Fig. 16: Diagrammatic representation of the square amplitude for the real gluon emission process $e^+e^- \to q\bar{q}g$.

To present the result of the computation we introduce the energy fractions:

$$x_i = \frac{2\,p_i \cdot Q}{Q^2} = \frac{2\,E_i}{Q}\ (\text{c.m. frame}) \quad \to \quad x_i > 0. \tag{37}$$

Energy conservation implies:

$$x_1 + x_2 + x_3 = \frac{2\sum_i p_i \cdot Q}{Q^2} = 2. \tag{38}$$

Furthermore, neglecting the mass of the quarks, we have that:

$$2\,p_1 \cdot p_3 = (p_1 + p_3)^2 = (Q - p_2)^2 = Q^2 - 2\,p_2 \cdot Q \tag{39}$$

that can also be written as:

$$2\,E_1 E_3 (1 - \cos\theta_{13}) = Q^2(1 - x_2). \tag{40}$$

This tells that $x_i < 1$ and that $x_2 \to 1$ in the limit $\theta_{13} \to 0$ with θ_{13} the angle between the three momenta of the outgoing quark with momentum p_1 and the gluon with momentum p_3. Including the phase space factors, in terms of the energy fractions, the real emission contribution to the cross section is given by:

$$\sigma^R = \int_0^1 dx_1 dx_2 dx_3 \delta(2 - x_1 - x_2 - x_3) |M_R(x_1, x_2, x_3)|^2 \tag{41}$$

and from the diagrams in Fig.(16) we have:

$$|M_R(x_1, x_2, x_3)|^2 = \sigma_0 C_F \frac{\alpha_S}{2\,\pi} \frac{x_1^2 + x_2^2}{(1 - x_1)(1 - x_2)} \tag{42}$$

So we see that, from the mathematical point of view, there are non-integrable divergences. This of course implies that the diagrams with the virtual emission will diverge in exactly the same way so to cancel the divergence of the real in the inclusive result. The collection of the divergences in Eqs.(41,42) is shown pictorially in Fig.(17).
Actually, from the running of the strong coupling and the experimental evidence, we already know that

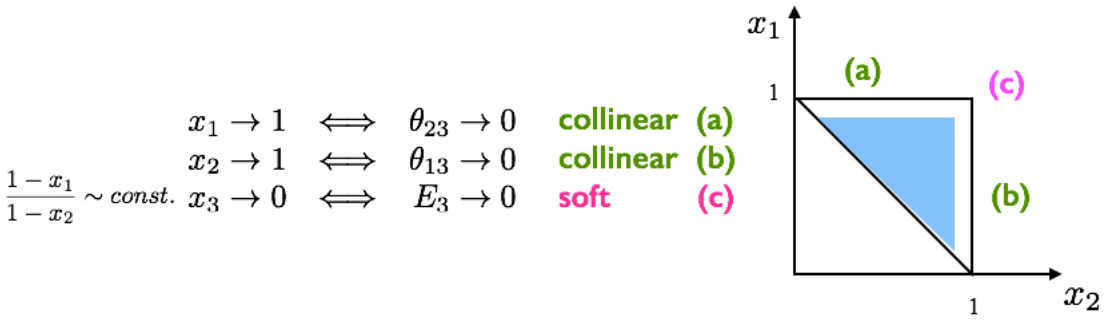


Fig. 17: Singular limits of the real matrix element squared and their representation in the 3-body Dalitz plot $x_1 x_2$. For massless partons, the available phase space is given by the upper right triangle. The singular limits corresponds to: the up ($x_1 = 1$) and right ($x_2 = 1$) edges, respectively corresponding to the configuration with the gluon becoming collinear to the quark and the anti-quark, and to the vertex (1,1) corresponding to a zero energy gluon (soft gluon limit).

there is a physical cut-off for the energy fraction, let's call it ε, given by the ratio of the typical hadron mass M_{had} and the hard scale of the process Q. Physically, here is indeed a change of regime for:

$$\varepsilon \sim \frac{M_{had}}{Q} \sim \frac{\Lambda_{QCD}}{Q} \tag{43}$$

Note that, however, even if the infinite might not be there, expressing the logarithmic dependence in term of the strong coupling, mathematically one has that:

$$\infty = \alpha_s(Q) \int_0^1 \frac{dx}{1-x} \quad \Rightarrow \quad \alpha_s(Q) \int_0^{1-\varepsilon} \frac{dx}{1-x} \sim \alpha_s(Q) \log \frac{1}{\varepsilon} \sim \alpha_s(Q) \frac{1}{\alpha_s(Q)}$$

from which one has (including more and more orders does not improve the situation):

$$\sigma \sim \sigma_0 \left(1 + \alpha_s(Q) \frac{1}{\alpha_s(Q)} + ...\right) \sim \sigma_0 (1 + 1 + 1 + ...). \tag{44}$$

Ultimately, then, the divergences that show up in the computation are intimately connected with the non perturbative regime of the theory and they are a characteristic feature of long distance (small ϵ limit) phenomena. It is useful to rewrite the real matrix element in Eq.(42) as:

$$\frac{x_1^2 + x_2^2}{(1-x_1)(1-x_2)} = \frac{1 + (1-x_3)^2}{(1-x_1)(1-x_2)} - 2 \tag{45}$$

and then use the relation:

$$\frac{1}{(1-x_1)(1-x_2)} = \frac{1}{x_3} \left(\frac{1}{1-x_1} + \frac{1}{1-x_2}\right) \tag{46}$$

so that the real matrix element, and in particular its most divergent part, gets the form:

$$d\sigma_R = \sigma_0 dW_{13} + \sigma_0 dW_{23} + \ finite\ terms \tag{47}$$

(and similarly for dW_{13}) with:

$$dW_{23} = \frac{dx_1}{1-x_1} dx_3 P_{qg}(x_3) = \frac{d\cos\theta_{23}}{1-\cos\theta_{23}} dx_3 P_{qg}(x_3) \tag{48}$$

and

$$P_{qg}(x_3) = C_F \frac{\alpha_s}{2\pi} \frac{1 + (1-x_3)^2}{x_3}. \tag{49}$$

The function $P_{qg}(x)$ is called splitting function and it is easy to show that

$$dW_{23} \xrightarrow[E_3\to 0]{\theta_{23}\to 0} \frac{d\theta_{23}^2}{\theta_{23}^2}\frac{dE_3}{E_3} \,. \tag{50}$$

In Eq.(48) we exploited the relation:

$$x_2x_3(1-\cos\theta_{23}) = 2(1-x_1) \implies x_1 = 1 - \frac{x_2x_3}{2}(1-\cos\theta_{23}). \tag{51}$$

Then, after performing the change of variable from x_1 to $\cos\theta_{23}$, the integration over the δ-function in Eq.(41) yields:

$$\begin{aligned}
&\int \frac{dx_1}{1-x_1}dx_2dx_3\delta(2-x_1-x_2-x_3) \\
&\qquad = \int \frac{d\cos\theta_{23}}{1-\cos\theta_{23}}dx_2dx_3\delta\left[1-\left(1-\frac{x_3}{2}(1-\cos\theta_{23})\right)x_2 - x_3\right] \\
&\qquad = \int_{-1}^{1}\frac{\cos\theta_{23}}{1-\cos\theta_{23}}\int_0^1 dx_3 \frac{1}{1-\frac{x_3}{2}(1-\cos\theta_{23})}
\end{aligned}$$

Now, it turns out that it is possible to combine the real and the virtual contribution together paying attention to the fact that the two contributions leave on different phase spaces. The result is as follows:

$$\begin{aligned}
\sigma_R + \sigma_V &= \sigma_0 \int_{-1}^{+1}\frac{d\cos\theta_{23}}{1-\cos\theta_{23}}\int_0^1 dx_3 P_{qg}(x_3)\left[\frac{1}{1-\frac{x_3(1-\cos\theta_{23})}{2}} - 1\right] \\
&\quad + (2\to 1) + finite\ terms
\end{aligned} \tag{52}$$

The first term in the squared bracket comes entirely from the real contribution, while the -1 is the contribution of the divergent part of the virtual term. Actually, knowing that the total is finite we could have predicted it. In the small x_3 (soft) and small θ_{23} (collinear) limits the whole squared parenthesis goes like $x_3\theta_{23}^2$, so regularizing the divergences of the other factors (see Eq. (50)). The finiteness of the inclusive result that we have seen in the previous section is however a property of the theory related directly to the unitarity that force the conservation of probabilities. In summary, the total of the virtual and the real contribution has to be finite (and their inclusive sum is just $\sigma_0\alpha_s/\pi$), but now we know something more: first, the virtual is divergent in every point of its (Born) phase space, second, the limiting configurations in which the real becomes divergent (extra radiated parton soft or collinear to another parton) are the ones in which the real and the virtual become indistinguishable, third, the two divergences cancels locally, ie for any and irrespective of the Born level partonic configuration. With these properties we can think to implement the local cancellation of the divergences and make predictions at differential level including the first order radiative corrections. This will have an impact on the kind of variables we can define in perturbation theory. Let's state it more carefully. In every computation that is not just the total cross section, the differential cross section is convoluted with a phase space function (F) that defines the physical quantity we want to compute (including experimental cuts):

$$\sigma_B = \int_m d\sigma_B \qquad\qquad d\sigma_B = d\phi^{(m)}|M_m^{tree}|^2F^{(m)}(p_i) \tag{53}$$

going at Next-to-Leading Order (NLO) we have:

$$d\sigma_V = d\phi^{(m)}|M_m^{loop}|^2F^{(m)}(p_i) \tag{54}$$

$$d\sigma_R = d\phi^{(m+1)}|M_{m+1}^{tree}|^2F^{(m+1)}(p_i) \tag{55}$$

so that:

$$\sigma_{NLO} - \sigma_B = \int_m d\sigma_V + \int_{m+1} d\sigma_R \tag{56}$$

In order for the InfraRed (IR) and Collinear (C) cancellation to take place also at differential level it is necessary and sufficient that the value of the observable we consider, F in the formulas above, is insensitive to a soft emission or to the situation in which the emitted parton is collinear to another parton. In formulas, one must have that for:

$$p_i \to 0 \qquad \text{or} \qquad p_i \parallel p_j \tag{57}$$

the function F is such that:

$$F^{m+1}(..., p_i, ..., p_j, ...) \simeq F^m(..., p_i + p_j, ...) \tag{58}$$

This condition guarantees the local cancellation of the singularities. As expected of course, the computed quantity cannot resolve long distance phenomena, but having cancelled the logarithmic enhancements from real and virtual, non perturbative physics will be power suppressed! The power of the suppression will depend on the physical observable. The implementation of the formula in Eq.(56) is however a highly non trivial task, but thanks to another fundamental property of QCD amplitudes this task has been strongly simplified. In the divergent limits in Eq.(57), the amplitudes factorize, in the sense that they can be approximated by the product of the (process dependent) amplitude with a final state parton less times a universal factor that represents the emission of the extra parton. Apart from colour and spin correlation factors (computable algorithmically), the amplitude with a parton less is just the leading order amplitude. The universality of the radiation factors can be exploited to write an auxiliary cross section to be subtracted from the real matrix element and so cancelling its divergences. One builds such a cross section combining the Born matrix element and the radiation factors, which bring the dependence on the radiation variables. To not alter the physical cross section one has to add back this subtraction and one can do so after having integrated the auxiliary cross section over the radiation degrees of freedom, obtaining a function that depends only on the Born phase space variables, as for the virtual contribution. So that, combining the two, one can also check the explicit cancellation of the divergences in the virtual matrix element. In formula, what is done in practice is to build two separately finite and integrable integrands as follows [10]:

$$\sigma_{NLO} - \sigma_B = \int_m \left[d\sigma_V + \int_1 d\sigma_{Aux} \right] F^{(m)}(p_i) + \int_{m+1} \left[d\sigma_R F^{(m+1)}(p_i) - d\sigma_{Aux} F^{(m)}(p_i) \right] \tag{59}$$

In the formula above we have made explicit the function that represents the measurement constraint, F. Note that the auxiliary cross section is in general a function of the whole real phase space variable. Nevertheless, it multiplies the observable function for the leading order kinematics $F^{(m)}$. Indeed, this must be the case in order not to alter the distributions at the differential level. In the language of Monte Carlo integration, we call a phase space point an event and we assign a weight to it accordingly to the integrand function. In particular, for the real events, the weight must be given by the real cross section $d\sigma_R F^{(m+1)}(p_i)$ only. The idea underlying Eq.(59) is that in correspondence of each real event we build a "counter-event" in the leading order phase space whose weight is given by the auxiliary cross section $d\sigma_A F^{(m)}(p_i)$. In the singular limits, event and counter-event coincide leading to the local cancellation of divergences and making the real integration finite. On the other hand, the extra weights in the leading order phase space due to the counter-events are exactly balanced by the integrated auxiliary cross section over the radiation degrees of freedom. This means that the procedure is effectively based on a mapping of the real phase space variables into the ones of the Born cross section plus (three) radiation variables, which allows one to pick a counter-event for each real event. Moreover, these radiation variables are those we integrate over to build the second piece in the first integrand in Eq.(59). Thanks to the universality of the factorization property and the possibility to write general mappings, the main ingredients of this

construction can be built once and for all and the auxiliary cross sections can be built algorithmically. Various subtraction formalisms have been proposed [11, 12] and they can be considered the prequel of the so called NLO revolution.

The emerging underlying theoretical picture of an high energy event is then as follows. Perturbation theory can be used to describe the hard process, ie the scattering process with the highest momentum transfer of the whole event. We learned from the running of the strong coupling that at very high energy scales it gets smaller and smaller, so that one can truncate the perturbative expansion at a given finite order to get a fully differential prediction for any InfraRed and Collinear safe observable. Then, perturbation theory can still be used to approximate the large number of subsequent emissions (almost soft and collinear) at lower and lower energies. Finally, when the original energy has been degraded and the emissions happen at energies close to the scale of the breakdown of the perturbation theory a new non-perturbative regime starts in which the interaction is really strong and binds quarks and gluons into hadrons. This last hadronization process, in which a multitude or shower of partons is packed into well separated and colourless hadrons, cannot certainly be described by perturbation theory and has to be modelled. This means that the hadronization process can be based on theoretically well motivated assumptions, but it also contains many parameters and ultimately no justification for their values. Nevertheless, in QCD the scaling of hadron multiplicity with the total energy of the e^+e^- annihilation can be computed with the only assumptions that hadron formation happens at low scale, of the order of 1GeV, and that perturbation theory is valid up to slightly above that scale. Relating hadron multiplicity to the average number of partons at the end of the parton shower cascade one can predict:

$$\langle n \rangle = \frac{\sqrt{96\pi}}{\beta\sqrt{\alpha_s(Q^2)}} + \left(\frac{1}{4} + 10\frac{n_f}{27\beta}\right)\log\alpha_s(Q^2) + \mathcal{O}(1). \tag{60}$$

Such a dependence on the collision energy of the number of charged particles produced, is in good agreement with data, as shown in Fig.(18).

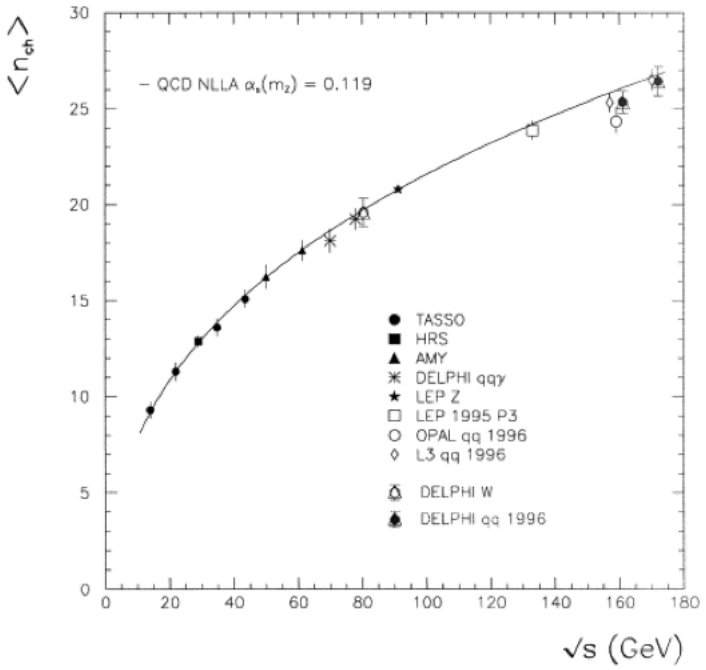


Fig. 18: Data-theory comparison for the average hadron multiplicity in a e^+e^- collision as function of the collider energy.

1.5 QCD phenomenology at differential level

1.5.1 Shape variables

The first example of an InfraRed and Collinear safe variable for e^+e^- annihilation into hadrons is the Thrust distribution. It is defined by the formula:

$$T = \max_n \frac{\sum_i |p_i \cdot n|}{\sum_i |p_i|} \tag{61}$$

where the p_i are the three momenta of the outgoing hadrons. Event by event, the value of the variable T is computed searching for the direction that maximize the ratio in Eq.(61). It is easy to convince yourself that a two jet like event, i.e. an event with just two hadrons or just two well collimated back to back sprays of hadrons, has $T = 1$, while an isotropic event ends up with $T = 1/2$. In turn, the value of this variable reflects the shape of the distribution of the hadrons in the final state. From this the name of shape variable for the Thrust. It is also easy to check that a soft or a collinear emission cannot alter significantly the Thrust value. Going at very high energy the strong coupling becomes smaller and smaller and so does the radiative correction to the average Thrust. The average Thrust is then predicted to tend to 1, approaching the leading order back-to-back configuration. This behaviour can be seen in Fig.(19) also in comparison with the experiments. In the fit reported in Figure(19) also power corrections

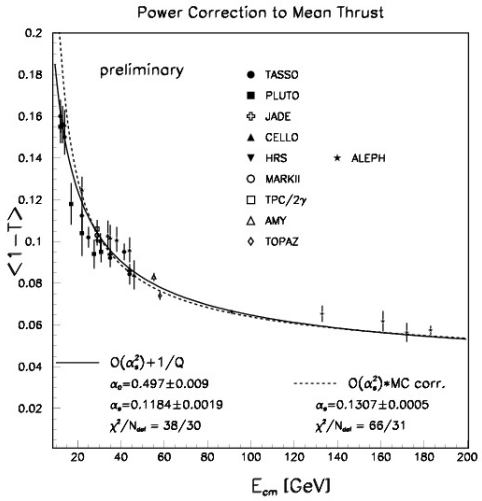


Fig. 19: Collection of average Thrust measurements in e^+e^- collisions at different collider energies. Actually, they are reported in terms of the variable $t = 1 - T$. As predicted by the QCD, the Thrust variable approaches 1 for increasing energies (so, correspondingly t tend to vanish). The solid line represents a fit which includes power correction of order Λ/Q to be compared to the dashed line fit (which does not include them).

to the (soft) hadronization process of order Λ/Q, are taken into account. Indeed, the emission of a pion with few hundred MeV transverse momentum, involving the strong coupling at such small scales, has probability 1. Using Eq.(61) the effect on the average Thrust measurement at LEP energies induced by such emission can be easily estimated to be of about 8%. This effect is parametrised fitting many data with different hadronization models. Going more differentially, perturbative QCD also predicts that the Thrust distribution must become more and more peaked in 1 with increasing energies as it is well verified experimentally (see Fig.(20), where the variable $t = 1 - T$ is represented). Thrust distribution can be used to measure the strong coupling constant. The fit range is the result of a balance among: size of the perturbative corrections, quality of data and size of the hadronization corrections that are usually estimated comparing predictions obtained using different Monte Carlo programs. A typical choice at

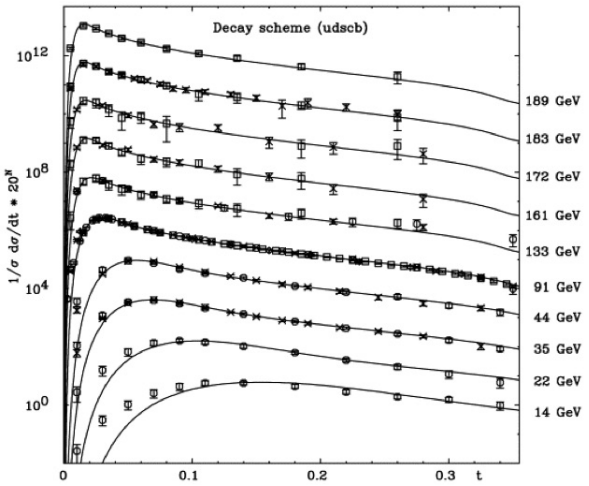


Fig. 20: Measurements of normalized Thrust distributions wrt to the variable $t = 1 - T$ for collisions at several energies. As predicted by QCD, the shape of the distribution becomes more peaked towards small values of t as the energy is increased.

LEP is shown in Fig.(21). The Thrust variable is only one example of a number of shape variables, more or less sensitive to the radiative corrections and so to the strong coupling constant. We will not discuss them further here and just point out that for the InfraRed and Collinear safety of the shape variables, it is the linearity wrt the particle momenta to play the main role, as in Eq.(61).

1.5.2 Jet cross sections

Experimentally, high energy e^+e^- collisions into hadrons produce for the large part two jet-like events, with two collimated back-to-back sprays of high energy hadrons, see Fig.(22). Then there is a lower number of events with three jets of particles, a much lower number of events where four jet structures can be identified and so on. This is in line with the expectations of perturbative QCD: we can imagine that the jet structures we see are in a one to one correspondence, and so they are the evolution, of the high energy partons that took part to the primary hard scattering. At high energy the coupling is relatively small and so the lower rate measured for increasing number of jets reflects the exponent of the coupling constant in the matrix elements squared for the production of such multi-parton final states. This is an heuristic argument, but for quantitative studies, one needs of a precise definition of a jet, which in turn is a prescription for grouping particles together. In particular, to the aim of comparing perturbative QCD predictions in terms of "partons" to measurements of collimated sprays of "paticles" we need a prescription that can be applied to both. Furthermore, for the theory prediction to be meaningful, one must be sure that the definition of a jet is inclusive enough wrt to the low energy particles in such way not to spoil the cancellation of long distance logarithmic divergences. In general, a jet algorithm must fulfill the following requirements: InfraRed (IR) and Collinear (C) safety, they should be simple to implement in both experimental analyses and theory predictions, and they should require small hadronization (non-perturbative) corrections. A way to proceed is to define:

1. a distance between particles, d_{ij};
2. a merging scheme to merge particles;

Then, one has to compute all the distances between particles (protojets) and merge the ones with minimum distance. This procedure is iterated until a fixed minimum distance among all protojets is reached.

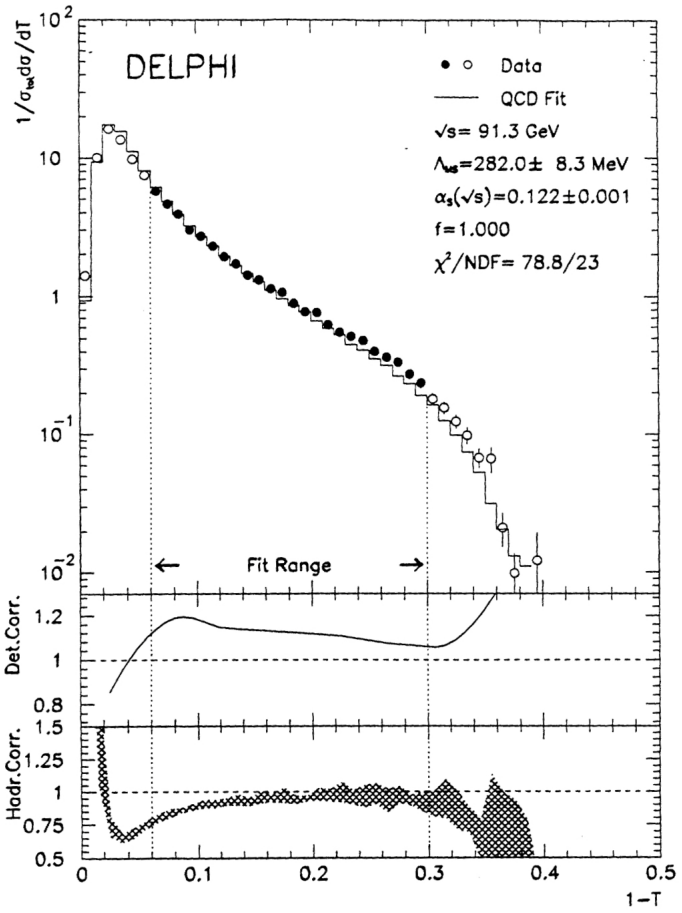


Figure 3:

Fig. 21: DELPHI dataset of measurements of the Thrust distribution to extract the strong coupling constant. The fit range is selected to minimize the impact of the detector and hadronization corrections.

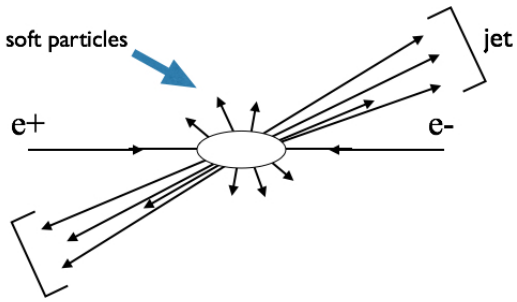


Fig. 22: Sketch of a two-jets event in e^+e^- annihilation.

Such a procedure is called an iterative clustering algorithm and unambiguously assigns particles to jets. As for IR and C safety, what matters is that the distance definition goes to zero if $E_i \to 0$ or $\theta_{ij} \to 0$. The number of jets will have, of course, a predicted dependence on the value of the fixed minumin distance (resolution) value.

The original JADE algorithm [13], for example, starts from a distance defined as $d_{ij} = 2E_iE_i(1 - \cos\theta_{ij})$ that for massless i, j particles represents the invariant mass squared of the pair. This is a dimensional variable, one can normalize it to the total energy of the event squared (Q^2) building the variable $y_{ij} = d_{ij}/Q^2$ and fix the resolution variable to some relatively small value y_{cut} so that the algorithm will proceed up to the situation in which $y_{ij} > y_{cut}$ for every pair of protojets i and j. Note, however, that with this algorithm soft particles turn out to be strongly correlated so that, even if they are at large angle, they are merged in the same jet and eventually can unnaturally produce a sort of sparse jet. With the kT algorithm [14], the distance measure is instead set to $d_{ij} = 2\,\mathrm{Min}\{E_i^2, E_j^2\}(1 - \cos\theta_{ij})$ that, in the small angle limit, tends to represent the transverse momentum of one particle wrt the direction of the other. Note that with this definition the distance d_{ij} is diagonal wrt particle energy. Soft particles are merged with the hard particle closest in angle (soft fragments are likely to be merged with their parent) avoiding unnatural assignments with creation of soft and wide angle jets. As for the merging scheme, sensible choices are simply to add the four momenta of the protojets (E-scheme) or first add and then rescale the space component of the sum to make again the new protojet massless (E0-scheme). Finally, once finished the assignment of all the particles to each jet one has to choose the way to assign the momentum to the final jets. A typical choice is to take the sum of the four momenta of the particles entering each jet. QCD predictions for jets rates, like their dependence on the algorithm and for each algorithm the increase of multijet rates by increasing the jet resolution ($y_{cut} \to 0$), have been nicely confirmed at LEP, for example. Furthermore, using a shower Monte Carlo it is possible to compare parton vs hadron jet rates to provide an estimation of non perturbative corrections. This, for example, motivates the introduction of the E0-scheme over the E-scheme, as fairly shown in the left panel of Fig.(23). In this figure the continuous line represents results at hadron level that reproduce well the data. Hadronization corrections turn out to be smaller for the kT algorithm (named D-scheme in the right panel Fig.(23)). Indeed, the kT algorithm is theoretically favoured also because the logarithmic unbalanced corrections that appear for small y_{cut}, can be resummed to all orders in perturbation theory. We conclude this section

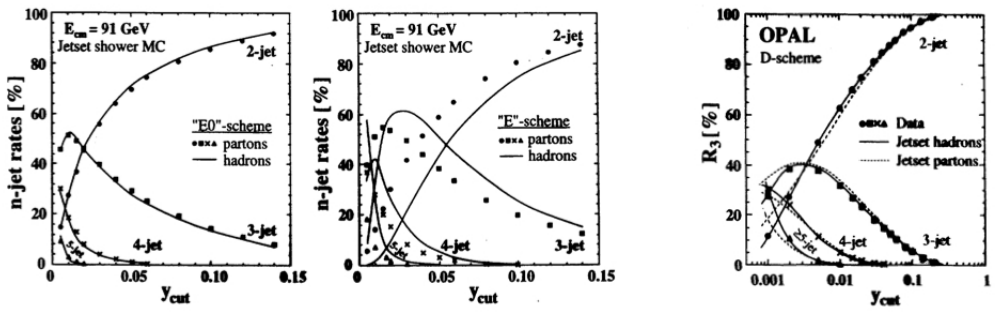


Fig. 23: Studies on jets reconstruction algorithms at LEP. Jets are reconstructed both at parton level (dots on left, dashed line on the right) and at hadron level (solid line). In the left panel, the E-scheme and the E0-scheme are compared: the latter has clearly better performances wrt to the hadronization corrections. In the right panel, the results for the kT algorithm.

showing in Fig.(24) [1] the nice combination of inclusive and differential measurements that prove the running of the strong coupling constant. In this figure there are also measurements extracted from deep inelastic lepton-hadron and hadron-hadron scattering, that will be the subject of the next sections, and the agreement of all of them with the predicted running is extremely good.

2 Hadrons in the initial state

The presence of hadrons in the initial state of a collision greatly complicates the picture. Nevertheless, for scattering processes characterized by a large momentum transfer, a description in terms of the parton

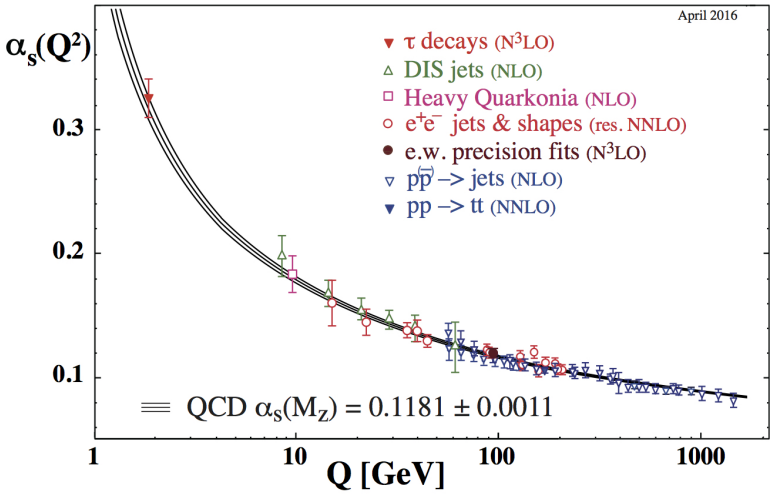


Fig. 24: Compilation of measurements of the strong coupling constant in a wide range of energies and for different physical processes. The solid lines represent the perturvative QCD prediction with its uncertainty bands.

model formulas is possible. Here "large" has to be understood as much larger than the characteristic hadronic binding energy. In this regime, it can be effectively thought that a single parton (quarks and gluons), and not the whole hadron, is taking part to the scattering process. Translating in a time-picture: if the scattering occurs on a time scale much smaller than the characteristic time of the processes that regulate the hadron dynamics, the system cannot respond to the interaction in a coherent way and a single parton is struck while the rest is frozen. So that, one assumes in general that a flux of composed particles (like atoms, nuclei or positronium bound states) is equivalent to a flux of their constituents carring just a fraction of the total moment with a distribution dictated by the intensity of the interaction among the constituents. Boosting the bound state particles along one direction produces a spread of the longitudinal component of the momentum while the transverse component remains the same as for the particle at rest (negligible after the boost). According to this picture, the parton formulas are convolutions of parton density functions (pdf's) and partonic scattering cross sections (with hat to distinguish them from particle cross sections). In case of hadron-hadron collision it reads:

$$\sigma_{H_1H_2}(p_1,p_2) = \sum_{i,j} \int dx_1 dx_2 f_i^{H_1}(x_1) f_j^{H_2}(x_2) \widehat{\sigma}_{ij}(x_1p_1, x_2p_2) \tag{62}$$

while for lepton-hadron collision one has:

$$\sigma_H(p,\gamma^*) = \sum_i \int dx f_i^H(x) \widehat{\sigma}_{i\gamma^*}(xp). \tag{63}$$

In summary, x in the formulas above is the fraction of the total momentum carried by the constituent partons and the variance of the fraction depends on the interaction strength. This scheme will work as far as we have large scattering angles (at small angle partons could scatter coherently) or, simply, as far as we can neglect the transverse momentum of the partons. With these formulas it is possible to make predictions at leading order for any scattering process involving hadrons in the initial state once the parton densities have been extracted from some measurement.

2.1 Naive parton model and Deep Inelastic Scattering

Let's start considering electron-hadron Deep Inelastic Scattering (DIS), namely the process

$$e^-(k) + P(p) \to e^-(k') + X$$

where we label with X any possible hadronic final state resulting from the fragmentation of the incoming proton as shown Fig.(25). Indeed, in a typical DIS experiment, one mainly measures the scattered leptons, inclusively wrt to the possible hadronic final states. The kinematics of the DIS scattering is

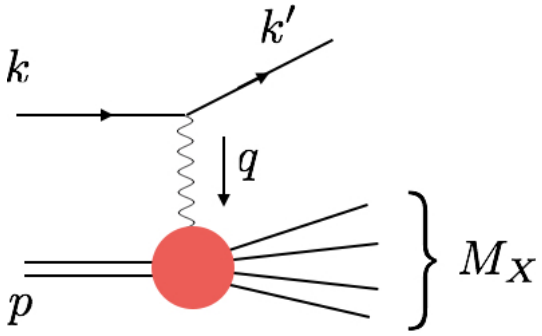


Fig. 25: Kinematics of electron-proton DIS scattering.

usually described in terms of the following variables: the square of the total energy in the CoM frame $S = (k+p)^2 = 2k\cdot p + M_P^2 \simeq 2k\cdot p$ (neglecting the proton mass M_P wrt to the energy of the incoming electron), the time-like momentum transfer $q = k - k'$, its virtuality taken with opposite sign to have the positive quantity $Q^2 = -q^2 > 0$, and the invariant scalar product $\nu = p\cdot q$. In terms of these variables, it is customary to introduce the dimensionless fractions:

$$x = \frac{Q^2}{2\nu} \qquad \text{and} \qquad y = \frac{q\cdot p}{k\cdot p} = \frac{2\,\nu}{S} \tag{64}$$

which are related to the energy transfer and the scattering angle of the outgoing electron in the lab frame. Then, one has that the invariant mass of the recoiling system X is given by:

$$M_X^2 = (p+q)^2 = Q^2\frac{1-x}{x}. \tag{65}$$

Note that, measuring the scattering angle and the energy of the outgoing electron, all the invariants defined above can be computed. What is observed in nature is that in the limit of large Q^2 and finite x the differential cross section scales with energy as:

$$\frac{d\sigma}{dxdy} \sim \frac{1}{Q^2} f(x,y) \tag{66}$$

with no hadronic scale involved!
Going to the parton language, one has to consider the elementary scattering process of a virtual photon off a massless quark as represented in Fig.(26). Again, we will use the hat to indicate partonic momenta,

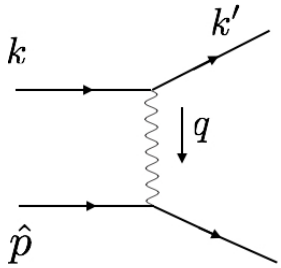


Fig. 26: Lowest order partonic elementary process contributing to the electron-proton DIS.

kinematic invariants and cross sections. The partonic variables are now:

$$\hat{s} = (k + \hat{p})^2 = 2\, k \cdot \hat{p} \qquad\qquad \hat{y} = \frac{q \cdot \hat{p}}{k \cdot \hat{p}} \tag{67}$$

and we impose the massless condition for the final state quark, $(q + \hat{p})^2 = 2\, q \cdot \hat{p} - Q^2 = 0$. It is easy to show that the partonic cross section is then:

$$\frac{d\hat{\sigma}_i}{d\hat{y}} = q_i^2 \frac{\hat{s}}{Q^4} 2\pi\alpha_{em}^2 \left[1 + (1 + \hat{y})^2\right] \tag{68}$$

where q_i are the partonic electric charge in units of the electron charge. Now, neglecting parton transverse momentum and assuming that the proton is a beam of partons carring each a certain fraction $\tilde{x}$ of the total hadron momentum, we have $\hat{p} = \tilde{x}\, p$, $\hat{y} = y$. From the massless quark condition ($2\, \tilde{x}\, p \cdot q - Q^2 = 0$) we find that the particle level and measurable $x = Q^2/2p \cdot q$ variable, coincides with the fraction of momentum $\tilde{x}$ carried by the scattered parton, so that we write:

$$\frac{d\sigma}{dxdy} = \sum_i f_i^H(x) \frac{d\hat{\sigma}_i}{d\hat{y}} \tag{69}$$

that inserting the expression for the partonic cross section gives:

$$\frac{d\sigma}{dydx} = \frac{2\,\pi\,\alpha_e^2\, S\, x}{Q^4} \left[1 + (1 + y)^2\right] \sum_i q_i^2 f_i^H(x). \tag{70}$$

The expression above predicts the mentioned scaling behaviour and also the full y dependence. It also shows that, of course, deep inelastic scattering is a good place to extract the parton densities $f_i^H(x)$, that are hadron specific. Indeed, the measured quantity is:

$$F_2^H(x) = x \sum_i q_i^2\, f_i^H(x) \tag{71}$$

Assuming isospin symmetry, i.e. that a neutron (n) is like a proton (p) with up and down quarks exchanged, we can combine data for electron scattering off protons and neutrons (deuterons) to extract $u(x)$ and $d(x)$:

$$F_2^p(x) = x \left(\frac{4}{9} u_p(x) + \frac{1}{9} d_p(x)\right) \tag{72}$$

$$F_2^n(x) = x \left(\frac{4}{9} u_n(x) + \frac{1}{9} d_n(x)\right) = x \left(\frac{4}{9} d_p(x) + \frac{1}{9} u_p(x)\right). \tag{73}$$

Note, however that now the interaction can create also quark pairs inside the hadron (very different from a non relativistic bound state like an hydrogen atom), so that indeed we are measuring $u(x) + \bar{u}(x)$ and $d(x) + \bar{d}(x)$, because the photon cannot distinguish the sign of the charge. Furthermore, each distribution can also diverge for very low momentum fraction, but still one has that sum rules related for example to the fermion number conservation:

$$\int_0^1 dx\, (u_p(x) - \bar{u}_p(x)) = 2 \qquad\qquad \int_0^1 dx\, (d_p(x) - \bar{d}_p(x)) = 1. \tag{74}$$

must hold. To get the difference among u and $\bar{u}$ we measure the structure functions related to charged current electroweak interactions like for neutrino DIS or $W^\pm$ production in hadronic collisions. As a result of this exercise to extract u, $\bar{u}$, d and $\bar{d}$ parton densities, one can look at the total momentum of the proton obtained summing over all the fractions:

$$P = \sum_i \int_0^1 dx\, x\, p_i \qquad p_i = u,\ \bar{u},\ d,\ \bar{d}. \tag{75}$$

This sum turns out to be smaller then 1 and it amounts to about $1/2$. We deduce once more that in the proton there are not only quarks but also something else that must interact rarely with photons (eventually indirectly), i.e. the gluons. To properly quantify the impact of the gluons in the description of the content of the proton probed at high momentum transfer we have to include radiative corrections.

2.2 Radiative corrections

The inclusion of the radiative corrections changes the parton model formulas given above. This is the reason why the formulas above are said the naive parton model formulas, while after including radiative corrections we have the so called improved parton model formulas. For hadron-hadron collisions for example we have:

$$\sigma_{H_1H_2}(p_1,p_2) = \sum_{i,j} \int dx_1 dx_2 f_i^{H_1}(x_1,\mu) f_j^{H_2}(x_2,\mu) \widehat{\sigma}_{ij}(x_1p_1, x_2p_2, \mu). \tag{76}$$

This formula has a similar form wrt the naive parton model formula, but the various factors have a different meaning. First of all, we introduce an arbitrary scale μ, which enters in all the factors in the rhs. In principle, we could distinguish among two different arbitary scales: the renormalization scale and another arbitrary scale to be called the factorization scale. However, as we will see in the following, in general it will be better to set them close to each other and in the range of the hard scale of the process, so we will use just one scale for the moment. The partonic cross section now has to be interpreted as a short distance cross section calculable in perturbation theory:

$$\hat{\sigma}_{ij}(x_1p_1, x_2p_2, \mu) = \sum_l \hat{\sigma}_{ij}^l(x_1p_1, x_2p_2, \mu)(\alpha_s(\mu))^l. \tag{77}$$

Furthermore, the dependence of the pdfs, $f_i^H(x,\mu)$, upon μ is mild and calculable with the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equation:

$$\frac{d}{d\log\mu^2} f_i^H(x,\mu) = \int_x^1 \frac{dz}{z} \sum_j P_{ij}(\alpha_s(\mu), z) f_j^H(x/z, \mu) \tag{78}$$

where the integral kernels P_{ij} are the Altarelli-Parisi splitting functions, also calculable in perturbation theory with a series expansion of the form:

$$P_{ij}(\alpha_s(\mu), z) = \frac{\alpha_s(\mu)}{2\,\pi} P_{ij}^0(z) + \left(\frac{\alpha_s(\mu)}{2\,\pi}\right)^2 P_{ij}^1(z) + ... \tag{79}$$

Finally, for small variations, the μ dependence in the pdfs and in the short distance cross section compensate. As stated above, the scale μ has to be chosen close to the hard scale of the process, this avoids the presence of large logarithmic corrections in the short distance cross section $\hat{\sigma}_{ij}$ (similarly to the renormalisation logarithm in e^+e^- annihilation). There will be, of course, divergences associated with radiation. We expect that soft and collinear divergences associated to gluon radiation off the final state will cancel (among the real and virtual contributions) as for e^+e^- annihilation into hadrons. For the initial state the situation is different: the collinear divergences of the real and virtual parts are associated to different kinematical configurations, in particular to final states with different invariant masses, and eventually do not cancel each other! This is a very different situation with respect to e^+e^- annihilation. Let's extract the most divergent part of the amplitude squared related to the initial emission as represented in Fig.(27).

We start from the kinematics. Imagine that the initial parton with momentum $\hat{p}$ is travelling along the z axis, then, to parametrize the divergences, we decompose the momenta as follows:

$$l = (1-x)\hat{p} + l_\perp + \xi\eta \tag{80}$$

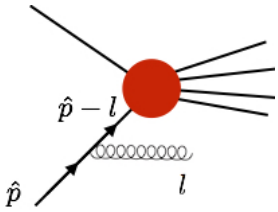


Fig. 27: Gluon radiation off an incoming quark line in a generic scattering process involving partons. In the strictly collinear limit, the amplitude exactly factorizes in a universal factor associated to the gluon emission process times the amplitude corresponding to the process without the radiation of the extra gluon.

with:

$$\hat{p} = (p_0, 0_\perp, p_0), \qquad \eta = (p_0, 0_\perp, -p_0), \qquad l_\perp = (0, \vec{l}_\perp, 0). \tag{81}$$

From the condition $l^2 = 0$ one has:

$$\xi = \frac{\vec{l}_\perp^{\,2}}{2\,\hat{p}\cdot\eta\,(1-x)} \tag{82}$$

and for the vanishing denominator of the propagator one has:

$$(\hat{p}-l)^2 = -2\,\hat{p}\cdot l = -2\,\hat{p}\cdot\eta\,\xi = -\frac{\vec{l}_\perp^{\,2}}{1-x}. \tag{83}$$

For the phase space factor one has:

$$\frac{d^3l}{2\,l^0\,(2\,\pi)^3} = \frac{d^2l_\perp}{2\,(2\,\pi)^3}\frac{dx}{1-x}. \tag{84}$$

The whole amplitude for the scattering process, that we generically indicate with A, is:

$$A = M\frac{\not{\hat{p}}-\not{l}}{(\hat{p}-l)^2}\not{\epsilon}u(\hat{p}). \tag{85}$$

In the above formula, we have stripped out the colour matrix t^A which will contribute with the colour factor C_F and we have denoted with M all the remaining part of the amplitude attached to the incoming quark line where the gluon emission occurs (which corresponds to the red blob and the other external legs in Fig.(27)). In the strict collinear limit $l_\perp \to 0$, the squared of the momentum $\hat{p}-l$ in the denominator is vanishing, exposing a singularity at the level of the amplitude squared which we would naively expect to behave as $d^2\vec{l}_\perp/\vec{l}_\perp^{\,4}$. However, this is not the case because when l becomes parallel to p also the scalar product of p and the gluon polarisation vector ϵ (that as we have seen is transverse for any physical gauge) vanishes. Indeed, the conservation of angular momentum provides a more general argument: there is no helicity flip in the gluon emission in a quark-quark-gluon vertex, so that the vectorial nature of the gluon provides suppression. With a bit of Dirac algebra one can show that:

$$(\not{\hat{p}}-\not{l})\not{\epsilon}u(\hat{p}) \simeq -\frac{2\,\epsilon\cdot l_\perp}{1-x}u(\hat{p}) + \not{\epsilon}\not{l}_\perp u(\hat{p}) = O(l_\perp). \tag{86}$$

The net effect is to make milder the behavior of the singularity which is actually logarithmic. Then, taking all together, the divergent part of the amplitude squared, the flux factor, the phase space factor, the colour factor, and integrating over the phase space, one has (from now on $l_\perp$ will denote the modulus of the bidimensional momentum $\vec{l}_\perp$) :

$$\sigma_R^{(1)} = C_F\frac{\alpha_s}{2\,\pi}\int\sigma^{(0)}(x\hat{p})\frac{1+x^2}{1-x}\frac{dl_\perp^2}{l_\perp^2}dx. \tag{87}$$

In this formula $\sigma^{(0)}$ represents the amplitude squared of the process after the initial gluon emission ($|M|^2$) with the momentum of the entering parton $\hat{p}$ reduced by a factor x by the emission, so that also the (invariant) center of mass energy of the collision is of course reduced by the same amount. This result actually is a consequence of the more general factorization property of the amplitudes in a gauge theory, which we mentioned in a previous section. The divergent part of the virtual correction can be recast in a similar form and all together the real and the virtual add to:

$$\sigma^{(1)} = C_F \frac{\alpha_s}{2\,\pi} \int \left(\sigma^{(0)}(x\hat{p}) - \sigma^{(0)}(\hat{p})\right) \frac{1+x^2}{1-x} \frac{dl_\perp^2}{l_\perp^2} dx. \tag{88}$$

So, we see that the soft divergence ($x \to 1$) cancels, while the collinear one ($l_\perp \to 0$) does not! Once again we can imagine that a lower cutoff scale λ of about 1 GeV regulates this divergence and that below that energy scale we enter another regime of the strong interaction. However, even in that case, our ability to make predictions in the perturbative regime is challenged because this construction implies a strong sensitivity to this IR cutoff λ. Note, that higher order corrections worsen the situation because the effective expansion parameter will be of order 1. Think of multiple (n) radiations over large $l_\perp$ gap, they will contribute to the event probability with a factor:

$$\left[\alpha_s(Q^2) \int_{\lambda^2}^{Q^2} \frac{dl_\perp^2}{l_\perp^2}\right]^n \tag{89}$$

where Q represents the hard scale of the process and each factor in the square parenthesis evaluates to:

$$\alpha_s(Q^2) \int_{\lambda^2}^{Q^2} \frac{dl_\perp^2}{l_\perp^2} = \alpha_s(Q^2) \log \frac{Q^2}{\lambda^2} \simeq \alpha_s(Q^2) \frac{1}{\alpha_s(Q^2)} = \mathcal{O}(1). \tag{90}$$

Regularisation of the collinear divergence will definitely not be enough, we will need of the resummation of the leading remaining effects. The problem is, indeed, solved with a renormalisation, similarly to the way we did the ultraviolet renormalisation for e^+e^- and it will work at the same time for all the observables receiving QCD corrections in hadron initiated collisions. In summary, the collinear not cancelled divergence is related to the vanishing of a denominator representing a massless propagator that is going on-shell. Furthermore the amplitude exactly factorises in the collinear limit into two amplitudes, one with the initial radiation and another one with the hard collision. The vanishing propagator implies the presence of a long distance physics phenomenon in the problem. Indeed, the collinear emission might well happen quite far from the hard interaction and it can be considered as another process among the ones that regulates the life of the hadron. This makes a big difference with respect to the naive parton model because, due to radiative corrections, the transverse momentum of the partons inside an hadron is not limited to Λ and in principle can be very large. We will now adsorb this collinear divergence into the definition of the parton distribution functions that make up the hadron. First, we rewrite Eq.(88) introducing an infrared cutoff λ to regulate the low $l_\perp$ emissions and we adopt the plus notation for the soft singularity getting a more compact form. Indeed, we observe that in Eq.(88), the function with a non-integrable singularity:

$$\frac{1+x^2}{1-x} \tag{91}$$

multiplies the difference among a regular function of the variable x and the same function evaluated in $x = 1$, in a such a way that the non integrable singularity is regularized. In analogy, a plus distribution version of the function in Eq.(91) is defined by the action on a generic function f as follows:

$$\int_0^1 \left(\frac{1+x^2}{1-x}\right)_+ f(x)dx = \int_0^1 \frac{1+x^2}{1-x} \left(f(x) - f(1)\right) \tag{92}$$

so that we can recast Eq.(88) in the form:

$$\sigma^{(1)} = \frac{\alpha_s}{2\,\pi} \log \frac{Q^2}{\lambda^2} \int \sigma^{(0)}(x\hat{p}) P^0_{qq}(x) dx. \tag{93}$$

The introduction of the plus distribution is useful also because the initial splitting functions are always the same for every process with any number of final state particles, in other words they are universal. In Eq.(93) we have also collected the colour factor and the plus distribution into the function P^0_{qq}. No surprise, this is once again the splitting function we have introduced in Eq.(49), while we were talking about e^+e^- into hadrons. Here there are two small differences: in Eq.(93) the x variable is the fraction of the original quark momentum carried by the quark after the splitting, while in Eq.(49) the fraction refers to the outgoing gluon and in fact to go from one expression to the other we have to exchange $x \to 1-x$, then the presence of the plus prescription that comes from the inclusion of the virtual correction. Note also that we have changed notation for the subscript, talking about the initial state we call P_{ab} the splitting probability to produce a parton of kind a from the decay of a parton of kind b, or even, to find a parton a inside a parton b, with fraction x of the momentum of parton b. Now we insert into the game the other (arbitrary) scale μ, that we mentioned at the beginning of the section, to split the logarithm. Neglecting terms of order α_s^2, we can write:

$$\sigma^{(0)}(\hat{p}) + \sigma^{(1)}(\hat{p}) = \int dx \left(\delta(1-x) + \frac{\alpha_s}{2\pi} \log \frac{\mu^2}{\lambda^2} P^0_{qq}(x) \right) \hat{\sigma}(x\hat{p}, \mu^2) \tag{94}$$

with:

$$\hat{\sigma}(x\hat{p}, \mu^2) = \hat{\sigma}^0(x\hat{p}) + \frac{\alpha_S}{2\pi} \log \frac{Q^2}{\mu^2} \int dz P^0_{qq}(z) \sigma^{(0)}(zx\hat{p}) \tag{95}$$

plus other finite terms. The scale μ is the factorisation scale and again, as we argue from Eq.(95), it will be better to choose it of the same order as Q to avoid large logarithmic corrections in $\hat{\sigma}(x\hat{p}, \mu^2)$. Now we convolute the corrected short distance cross section with the parton density getting:

$$\sigma(p) = \int dy dx f_q(y) \left(\delta(1-x) + \frac{\alpha_S}{2\pi} \log \frac{\mu^2}{\lambda^2} P^0_{qq}(x) \right) \hat{\sigma}(xyp, \mu^2). \tag{96}$$

Finally, we make the last manipulation introducing a delta function to get:

$$\sigma(p) = \int dz \widetilde{f}_q(z, \mu^2) \hat{\sigma}(zp, \mu^2) \tag{97}$$

with:

$$\widetilde{f}_q(z, \mu^2) = \int dy dx f_q(y) \left(\delta(1-x) + \frac{\alpha_S}{2\pi} \log \frac{\mu^2}{\lambda^2} P^0_{qq}(x). \right) \delta(z - xy) \tag{98}$$

The last step has been the adsorption of the large logarithmic correction into a redefinition of the parton density. From the universality of the splitting probabilities, it can be shown that this parton density redefinition does not depend upon the specific hard process, and so it's universal! Furthermore, there is a variety of arguments showing that the above construction holds to all orders in perturbation theory (factorisation theorem). In turn, we have to consider a parton as having a structure that depends upon the scale at which we are probing it ($\mu \sim Q$).

Nevertheless, we are now left with these large corrections that depend upon unknown low scale dynamics. The key observation here is that although low scale dynamics of parton density functions can only be measured, their scale dependence is predictable in perturbation theory. Indeed, neglecting terms of order α_S^2 we have that:

$$\frac{d}{d \log \mu^2} \widetilde{f}_q(z, \mu^2) = \frac{\alpha_S}{2\pi} \int dy dx \widetilde{f}_q(y, \mu^2) P^0_{qq}(x) \delta(z - xy) + \mathcal{O}(\alpha_S^2) \tag{99}$$

that is the Altarelli-Parisi, or Dokshitzer-Gribov-Lipatov-Altarelli-Parisi equation [15]. Once the pdf's are measured in a certain set of measurements performed at a given energy scale, they can be used as the initial conditions for the Altarelli-Parisi equations. Then, the solution to these equations, corresponding to the effective resummation of the logarithms associated to multiple collinear emissions, provides the pdf's to all other scales that can be used to make predictions for processes at different energy scales.

2.3 Altarelli-Parisi splitting probabilities and evolution equations

Indeed, the program outlined at the end of the last section needs of the knowledge of the splitting functions for all the possible elementary processes and of the solution of the following system of integro-differential equations:

$$\begin{aligned} \frac{d}{d\log\mu^2} f_q(x,\mu^2) &= P_{qq}\otimes f_q + P_{q\bar{q}}\otimes f_{\bar{q}} + P_{qg}\otimes f_g \\ \frac{d}{d\log\mu^2} f_g(x,\mu^2) &= P_{gq}\otimes f_q + P_{g\bar{q}}\otimes f_{\bar{q}} + P_{gg}\otimes f_g \end{aligned} \tag{100}$$

where we have used the $\otimes$ symbol to represent the convolution integral:

$$f\otimes g \equiv \int dydz f(y)g(z)\delta(x-yx) = \int_x^1 \frac{dz}{z} f\left(\frac{x}{z}\right) g(z). \tag{101}$$

The computation of the lowest order splitting functions is not difficult and can be done following the steps of the previous paragraph also for the other splitting processes. The results are given in Fig.(28). These

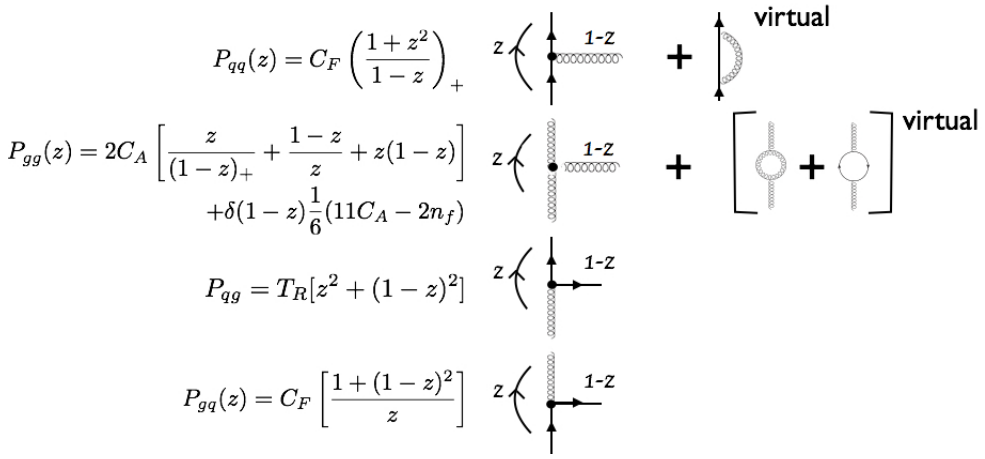


Fig. 28: Lowest order Altarelli-Parisi splitting kernels.

functions fulfil a number of properties: first, P_{qq} and P_{gg} exhibit the soft singularity at $z=1$, while of course P_{gq} has it for $z=0$, all of them are positive definite for $z<1$, the real parts are connected by final state parton exchange, it's easy to prove also the crossing symmetry $z\to 1/z$ keeping into proper account the number of states while building the colour/spin averages. Furthermore, the evolution does not spoil the sum rules of course. Considering the proton, we report again the flavour sum rules:

$$\int_0^1 \left[f_u(x,\mu^2) - f_{\bar{u}}(x,\mu^2)\right] = 2 \tag{102}$$

$$\int_0^1 \left[f_d(x,\mu^2) - f_{\bar{d}}(x,\mu^2)\right] = 1. \tag{103}$$

further, from momentum conservation we now must have:

$$\sum_a \int_0^1 dx\, x\, f_a(x,\mu^2) = 1. \tag{104}$$

Differentiating with respect to the scale μ one obtains the following conditions respectively:

$$\int_0^1 [P_{qa}(x) - P_{\bar{q}a}(x)] = 0 \qquad \forall\, a \tag{105}$$

$$\sum_a \int_0^1 dx\, x\, P_{ab}(x) = 0 \qquad \forall\, b \tag{106}$$

and the splitting probability functions quoted in Fig.(28) satisfy them. As a power series in the strong coupling, the splitting functions are known up tp Next-to-Next to Leading Order (NNLO). They contain terms proportional to $\log \mu^2$ but also to $\log 1/x$ and $\log(1-x)$. The leading order DGLAP evolution equation sums up the $(\alpha_S \log \mu^2)^n$ contributions, the Next to Leading order (NLO) sums the $\alpha_S(\alpha_S \log \mu^2)^{n-1}$ terms and so on. Small x resummation has also been performed predicting a power low behaviour for the pdf's. In general the precision of the experimental data demands that at least NLO (and preferably NNLO) DGLAP evolution has to be used in comparisons between theory and experiment.

Finding a stable solution of the system of Eq's(100) is a non trivial numerical task that is usually addressed with the Runge-Kutta method. Nevertheless, for extreme situations we can predict the leading behaviour of the evolution. Let's consider the large x limit first, $x \to 1$. In this case the dominant splittings are P_{qq} and P_{gg} so that the equations decouple and the leading terms in these two splitting functions are the two plus distributions that can be both represented as:

$$P_{aa} \simeq \frac{2\, C_a}{(1-z)_+} \tag{107}$$

with C_a equal to $C_F(C_A)$ for the quark q(gluon g). This is the limit of soft gluon radiation. Starting from the known pdf's at the scale Q_0, the evolution takes the form:

$$f_a(x,Q^2) \simeq f_a(x,Q_0^2) \exp\left\{ \int_x^1 dz \frac{2\, C_a}{(1-z)_+} \int_{Q_0^2}^{Q^2} \frac{dq^2}{q^2} \frac{\alpha_S(q^2)}{2\,\pi} \right\}. \tag{108}$$

From the definition of the plus distribution one has that:

$$\int_x^2 dz \frac{1}{(1-z)_+} = -\int_0^x \frac{dz}{1-z} = \log(1-x) \tag{109}$$

and inserting the running coupling we find:

$$f_a(x,Q^2) \simeq f_a(x,Q_0^2)(1-x)^{p_a} \qquad \text{with} \qquad p_a = \frac{C_a}{\pi\, b_0} \log \frac{\alpha_S(Q_0^2)}{\alpha_S(Q^2)}. \tag{110}$$

Then, we see that in the large x limit, the pdf's with our approximations vanish with a power-law behavior. Increasing the scale the exponent slightly increases, so that the higher is the scale the flatter is the behaviour of the distribution. In the small x limit instead, only the second term in P_{gg} matters. The inclusion of the splitting described by P_{gq} does not catch the leading singularity because it cannot give rise to a chain of all enhanced contributions. This is the limit driven by multiple soft gluon exchange so that only the equation for the gluons is relevant, with:

$$P_{gg} \simeq \frac{2\, C_A}{z}. \tag{111}$$

In this case the solution takes the form:

$$xf_g(x,Q^2) \simeq xf_g(x,Q_0^2) + \exp\left\{\sqrt{\frac{2\,C_a}{\pi\,b_0}\log\frac{\alpha_S(Q_0^2)}{\alpha_S(Q^2)}\log\frac{1}{x}}\right\}. \tag{112}$$

Inspecting the exponent and knowing the running coupling, this expression predicts a steepness increasing with Q^2 at small x.

To compare with phenomenology, we go back to the proton F_2 structure function and now we move from the naive parton model prediction:

$$F_2(x) = x\sum_i q_i^2\, f_i(x) \tag{113}$$

to the one including the leading radiative corrections (with $\mu^2 = Q^2$) resumed in the pdf evolution and get:

$$F_2(x,Q^2) = x\sum_i q_i^2\, f_i(x,Q^2). \tag{114}$$

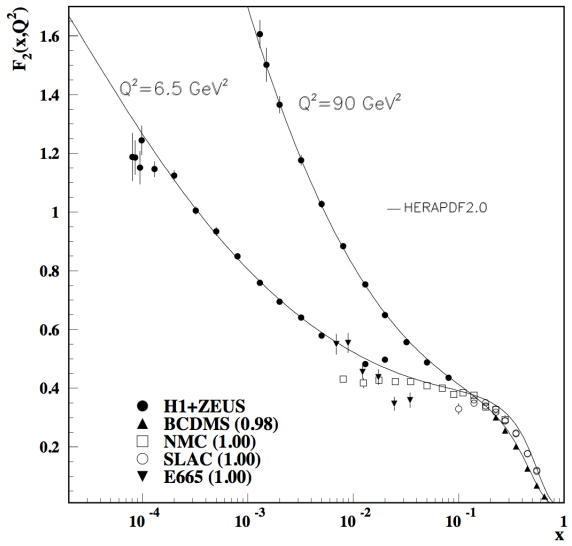


Fig. 29: Fit of the proton F_2 structure function as function of x for two different scale $Q^2 = 6.5\,\mathrm{GeV}^2$ and $Q^2 = 90\,\mathrm{GeV}^2$. As predicted by the DGLAP evolution, with increasing energy, the pfds become steeper at small x and flatter at large x.

In Fig.(29) [1] we can see the effects of the scale evolution mentioned above: the faster descent at high x and the greater steepness at low x. The scaling violations together with the QCD analysis is also shown in Fig.(30). The best fit of the pdf's is performed combining a large amount of data for all kind of available collisions and at very different energies. In Fig.(31) a sketch of the kinematic domain of the available data used in pdf fits is shown, while in Fig.(32) a solution obtained by the NNPDF [16] collaboration is shown for two scale values, similar results are obtained from other collaborations (see for example ABM [17], JR [18], MSTW [19], MMHT [20] and other collaborations).

3 Hadronic collisions and jets

3.1 Minimum bias

Proton-proton and proton-antiproton colliders are mainly discovery machines, see Fig.(33). In these

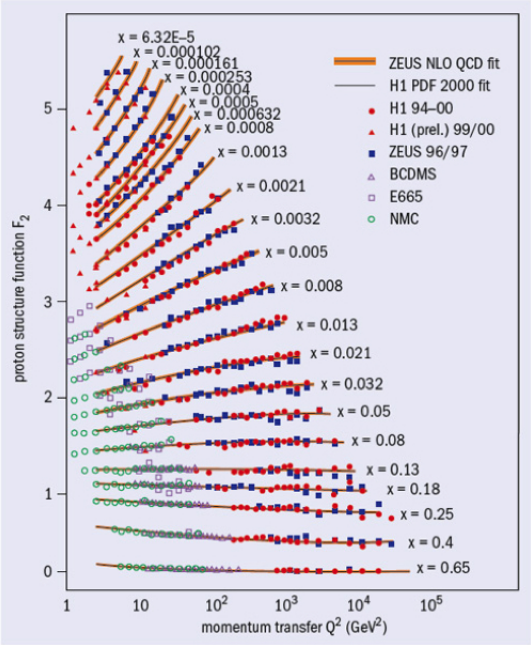


Fig. 30: Fits of the proton F_2 structure function as a function of the momentum transfer Q^2 for different values of the momentum fraction x. The simple scaling low with Q^2 at fixed x predicted by the parton model breaks down due to radiative corrections that well explain the measured scaling violation.

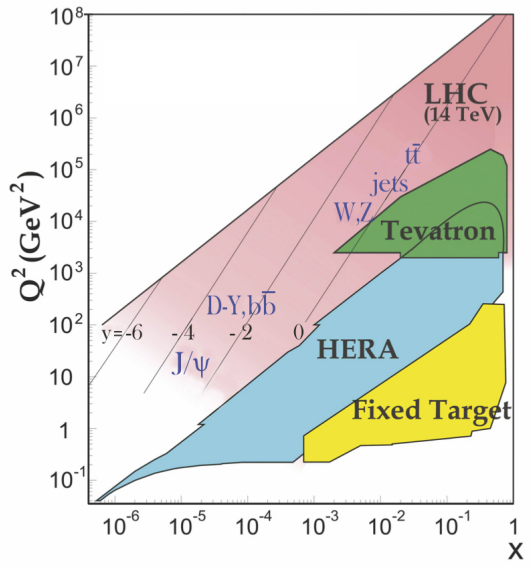


Fig. 31: Available data sets for the pdf fit in the plane $x - Q^2$.

experiments, QCD is ubiquitous, and this situation of course allows one to make many tests of the theory. Furthermore, a good knowledge of QCD effects is of course essential to establish signal and background rates. The variables used to describe hadron production in hadron-hadron collisions are represented in Fig.(34). The transverse plane is the plane orthogonal to the beams and the azimuthal angle is the azimuth around the beam direction. The transverse momentum is the projection of momentum on the transverse plane, $k_\perp = k|\sin\theta|$ and the transverse energy is the quantity $E_T = E|\sin\theta|$. Sometimes the transverse

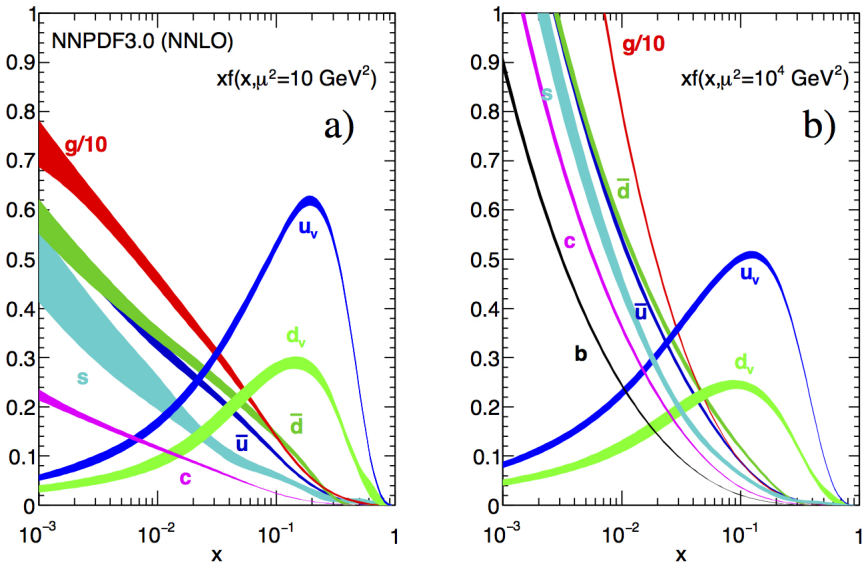


Fig. 32: Examples of pdf fits including DGLAP evolution at NNLO by the NNPDF collaborations. On the left, the pdfs at a scale $Q^2 = 10\,\text{GeV}^2$ while on the right $Q^2 = 10^4\,\text{GeV}^2$.

Accelerator	SppS	TEVATRON	LHC
Energy	200÷900GeV	1.96TeV	13TeV
Discovery	W/Z,Jets	top	Higgs

Fig. 33: Summary of the main discoveries achieved at the proton-antiproton (SppS and Tevraton) and proton-proton colliders (LHC).

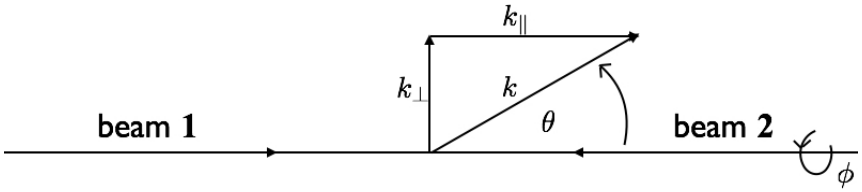


Fig. 34: Representation of the kinematical variables used in hadron-hadron collisions.

mass $m_T = \sqrt{k_\perp + m^2}$ is also used. The rapidity is defined by the formula:

$$y = \frac{1}{2}\log\frac{k^0 + k_\parallel}{k^0 - k_\parallel} \tag{115}$$

where k^0 is the energy of the detected particle. Under a boost in the beam direction it gets an additive contribution related to the boost velocity, so that rapidity differences are invariant under such boosts. This is particularly useful for the description of hadron-hadron collisions where there is always a boost among the hadronic and partonic cms. For massless particles the rapidity reduces to:

$$y = \frac{1}{2}\log\frac{1+\cos\theta}{1-\cos\theta} = -\log\tan\frac{\theta}{2}. \tag{116}$$

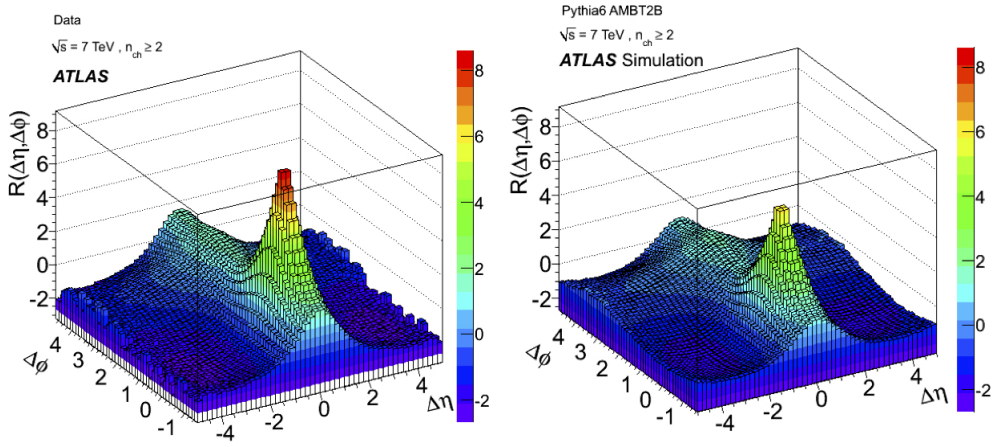


Fig. 35: Study of the angular correlations among hadrons in the Minimum Bias (main cuts: $k_\perp > 100\,\text{MeV}$ and $|\eta| < 2.5$) by the ATLAS collaboration. The data distribution (on the left) is compared to a simulated sample using the PYTHIA events generator (on the right). As stated in the main text, the simulation programs do not reproduce well the peculiar features of the data.

This variable refers directly to the scattering angle in hadronic cms frame and, as such, can be defined also for massive particles. In this case it is called pseudorapidity (η). Note however that η differences are not anymore boost invariant so that it is not immediate (as for the case of rapidity) to compare pseudorapidity distributions, for example among the results of hadron collisions in fixed target experiment with collider results, and to this aim one has to reinsert back the mass of the measured particles. As we have seen in the first section, the cross section for hadronic collisions reflects the presence of a typical hadronic scale:

$$\sigma \simeq \frac{1}{(\text{Few hundred MeV})^2} \tag{117}$$

and it increases very slowly with the center of mass energy. These features are not computable in QCD although this behaviour is consistent with QCD. Consider quark-quark scattering down to a minimum exchanged virtuality Q^2, one has that:

$$\frac{d\sigma}{dQ^2} \sim \frac{1}{Q^4} \Longrightarrow \int_{Q^2_{min}}^{Q^2_{max}} \frac{d\sigma}{dQ^2} dQ^2 \sim \frac{1}{Q^2_{min}} \tag{118}$$

so that, the typical scale of the non perturbative phenomena, naturally appears and QCD is in line with the fact that these kind of low energy interactions dominate. The bulk of the scattering events is then soft and rarely we assist to a collision characterized by a relatively large momentum transfer. We collectively call the soft collisions Minimum Bias (MB). A typical collision event at the LHC produces on average $80 \sim 100$ charged particles with huge fluctuations. This number grows only (a bit more then) logarithmically with the cms energy of the colliding protons. The transverse momentum distribution for a species of hadron with mass m has an exponential descent behavior ($\sim \exp\left(-a\sqrt{m^2 + k_\perp^2}\right)$ with an average value of $\langle k_\perp \rangle \sim 600\,\text{MeV}$ which, again, grows slowly with energy. As for rapidity distributions, they are rather flat with a total of $5 \sim 6$ particles per unit of central rapidity. Keep in mind that for a particle of mass $\sim 1\,\text{GeV}$ produced at $E_{CM} = 13\,\text{TeV}$ the maximum absolute rapidity is about 9. Again, all these features are non computable in QCD, but are compatible with QCD. To understand this one can just start from heavy quark pair production with a mass $m \gg \Lambda$. Reducing the value of the mass, one gets a broadening of the rapidity distribution and the result that the typical transverse momentum is of the order of the mass of the produced object. The features of the minimum bias constitute the first study performed at hadron colliders, which starts considering the angular correlation among the

produced hadrons in a sufficiently large number of collisions, collecting particles up to very low values of the transverse momentum. In Fig.(35) [21], the left panel shows the result of a study by the ATLAS collaboration obtained collecting hadrons with $k_\perp > 100\,\text{MeV}$ and $|\eta| < 2.5$ (plus some other less relevant cuts). We can clearly see: 1. the emergence of a jet like structure at $\Delta\eta = \Delta\phi = 0$, 2. the emergence of the recoil of one particle against the other as a back to back jet like structure at $\Delta\eta = \pi$ for all $\Delta\phi$ and the 3. decay of resonances or other colour structures at $\Delta\eta$ of about 2 and for all $\Delta\phi$. Nevertheless, these three features are not well reproduced by simulation programs (see for example the right panel in Fig.(35)) and more investigation is ongoing on this point.

3.2 Jets

For jets there is now a very different situation with respect to e^+e^- annihilations, where almost all events are back to back. From high energy partonic collisions we expect something similar, although now the higher the energy of the jets, the smaller the fraction of events. To identify jets something has to be done to distinguish them from the fluctuations of the minimum bias. Indeed, trigger is crucial to make discoveries! Think about the discovery of the Z boson at the UA2 experiment with proton-antiproton collisions at 630 GeV cm energy. In that case one has:

$$\frac{\sigma(Z)}{\sigma_{TOT}} \cdot Br(Z \to l^+l^-) \sim \frac{\frac{1}{3}\left(\frac{1}{300\text{GeV}}\right)^2}{\left(\frac{1}{300\text{MeV}}\right)^2} \cdot 0.1 \approx \frac{1}{3}10^{-7}. \tag{119}$$

Considering that one year has about $3 \cdot 10^7$ seconds, to have 1 Z boson decaying to charged leptons per year, you need 1 proton collision per second, while to claim the discovery you need more, O(1000)! This means that you need to register events at a rate of the order of few thousands per second. This is not a problem today, with the modern systems of data acquisition, thirty or forty years ago it was different. Furthermore, as stated in the previous section, most of the collisions involved an overwhelming rate of soft QCD reactions and particles which we are not interested in and that can represent a real challenge for the acquisition electronics in terms of memory resources and time response. Then, what is needed is an efficient system to choose the events to store (the trigger). As a general guideline, one stores events with a final state characterized by particles with large transverse momentum, that is for sure the signal of a short distance interaction. One could think to trigger events on the base of the presence of very high energetic particles in an event, but that is not a very good idea because minimum bias events can have very high energetic particles at small angles. It is better to use a "transverse energy trigger" (energy suppressed by the angle):

$$E_T = \sum_i E_i |\sin\theta_i| \tag{120}$$

But still, minimum bias events at UA2 counted ~ 25 charged as well neutral hadrons, so having an average transverse energy $\langle E_T \rangle \approx 50 \cdot 500\,\text{MeV} \approx 25\,\text{GeV}$. To stay away from minimum bias events the experiment considered events with $E_T > 70\,\text{GeV}$ finding that all events are as in Fig.(36), i.e. jet-like! Most of the events have two jets, firmly establishing that jets in hadronic collisions are a property of nature. From the definitions of rapidity and transverse momentum it is easy to show that measuring y_1, y_2 and $p_\perp$ we fix the values of the fractions x_1 and x_2. We can make theory predictions using the improved parton model formula for two jet differential cross section:

$$\frac{d^3\sigma}{dy_1 dy_2 dp_\perp^2} = \frac{1}{16\pi S^2} \sum_{i,j,k,l} \frac{f_i(x_1,\mu)}{x_1} \frac{f_j(x_2,\mu)}{x_2} \bar{\sum} |M_{i,j,k,l}|^2 \frac{1}{1+\delta_{kl}}. \tag{121}$$

and parton densities extracted from DIS measurements. The formula above predicts that jets are back to back in azimuth and it also predicts the jets angular distribution in the cm frame of the two jets (or equivalently the $p_\perp$ dependence of the jets). This differential prediction is nicely verified experimentally,

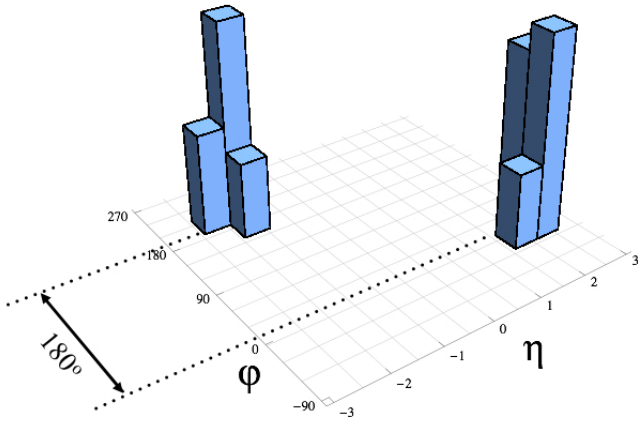


Fig. 36: Two jets event at UA2 experiment as emerging after imposing the hard cut on the transverse energy $E_T > 70\,\text{GeV}$ to separate the hard scattering events from the minimum bias.

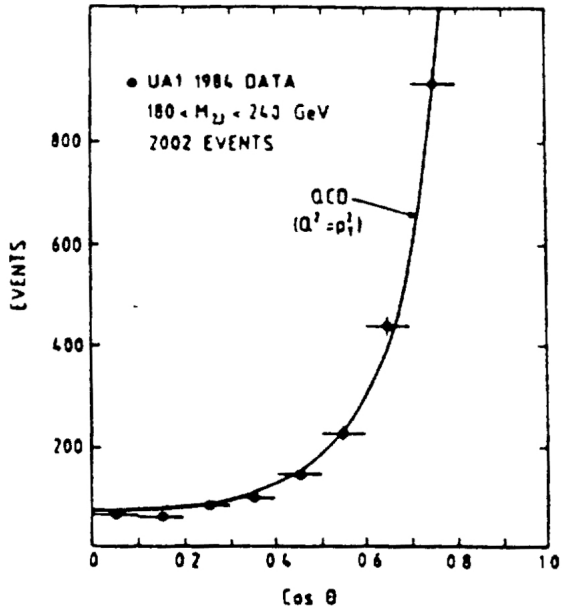


Fig. 37: Measurement of the jets scattering angle in the jets cm frame by the UA1 experiment. The solid line is the QCD prediction given by the improved parton model formula.

as it can be seen in Fig.(37). Furthermore, note that as a good first approximation, the partonic scattering cross sections stay with colour factors dictated as shown in Fig.(38). So that, the two jet cross section can be written in terms of a generalized structure function F as:

$$\frac{d^3\sigma}{dx_1 dx_2 d\cos\theta} = \frac{F(x_1)}{x_1}\frac{F(x_2)}{x_2}\frac{d\hat{\sigma}_{gg\to gg}}{d\cos\theta} \tag{122}$$

with:

$$F(x) = f_g(x) + \frac{4}{9}\sum_q \left[f_q(x) + f_{\bar{q}}(x)\right]. \tag{123}$$

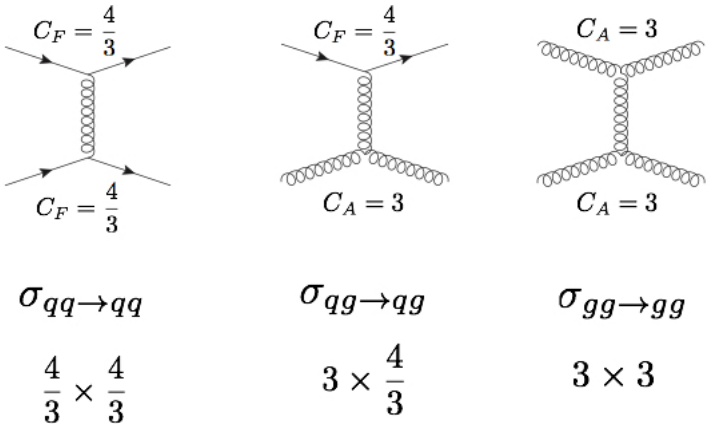


Fig. 38: Main (lowest order) partonic subprocesses contributing to the jets production at hadron collision and their corresponding colour structure.

In Fig.(39), we can see the comparison among theory and experiment. Note the necessity to include the gluon pdf to get agreement.

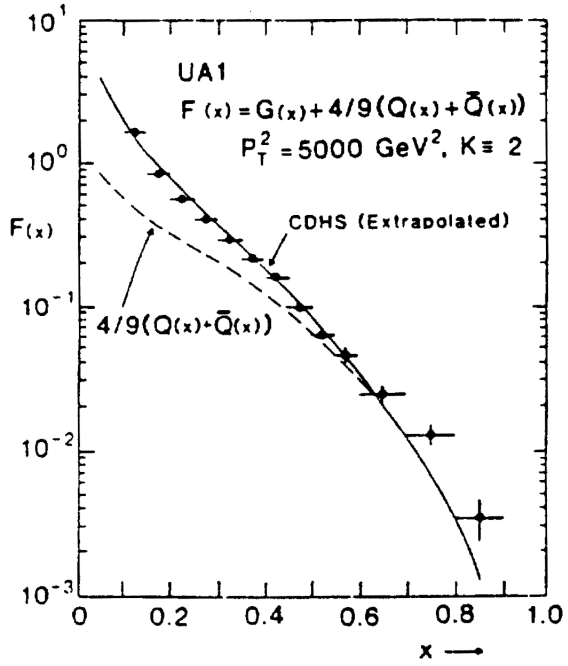


Fig. 39: Measurement of the jet generalized structure fuction $F(x)$ as function of the momentum fraction x by the UA91 experiment and comparison with the theory prediction. The dashed line corresponds to the QCD excluding the gluon contribution. The inclusion of the gluon pdf (solid line) is required to have agreement with the data.

3.3 Jets algorithms

By the way, it is only when we go to higher orders in perturbation theory that we start getting more solid predictions for the normalization of the cross sections with a reduced dependence from the unavoidable but unphysical scales (renormalization and factorization scales). Such predictions allow to perform more reliable tests of QCD comparing theory and experimental measurements of jet differen-

tial distributions [22]. However, not all the measurements that can be done at an hadron collider can be compared to QCD perturbation theory. QCD can provide perturbative predictions only for InfraRed and Collinear safe variables. For jet physics, as we will see in a moment, this means that care must be given for theoretically viable jet definitions. At very high energies like at the Tevatron or the LHC, for the large part of the interesting events the situation is not as clear as the one in Fig.(36). Looking at the η, ϕ plane, one often has a situation more like the one in Fig.(40). with the energies of the particles

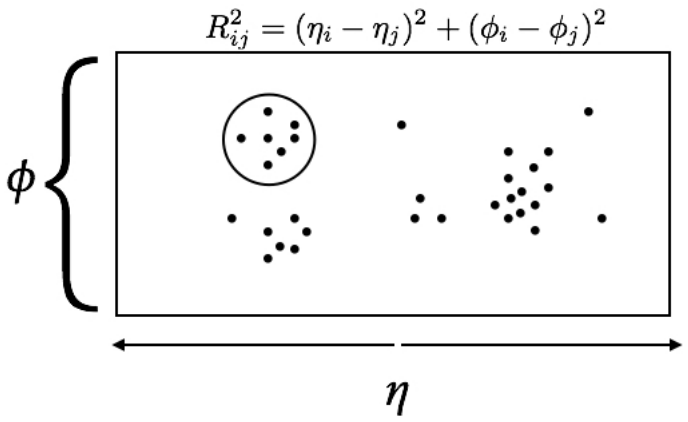


Fig. 40: Distribution of the reconstructed particles in the η, ϕ plane for a typical collision event at the Tevatron or the LHC. There are sparse groups of particles and the assignment particle-jets is not unambiguously clear.

(or small groups of particles in a calorimeter cell) that are very different from each other. The basic idea is always that, in very high energy collisions, jets can be reconducted to the original partons that came out from the hard scattering. Whenever shower and (non perturbative) hadronization corrections are expected to be small, the same algorithm can be logically applied to both partonic and particle level data. It could happen that for experimentally well motivated jet definitions, through which it is possible to make very good measurements, it is not possible to make higher order predictions of jet observables to compare with, without introducing sensitivity to long distance phenomena. This is indeed the case for several cone algorithms. A cone algorithm is one in which a cone is drawn around the jet axis and the particles contained in that cone are collected and assigned to the jet. The crucial point is how to fix the jet axis. Consider for example for the so-called highest-E_T-seed algorithm. Once you fix the cone radius R, this works as follows: 1. select the highest E_T particle as the jet axis, 2. build the cone and remove the particles in it, 3. restart from 1. One can perform measurements of the jet activity in hadron-hadron collisions using this algorithm, but it is not hard to see that the jet number is a Collinear unsafe observable and so the perturbative prediction will get at each order logarithmic enhanced corrections that depend on the low energy dynamic. The problem can be visualized thinking to a system of three partons with transverse energies such that $E_{T1} > E_{T2} > E_{T3}$ and $E_{T1} < 2\,E_{T2}$. Furthermore, assume that the distances are such that $d_{21}, d_{31} < R$. This configuration gives rise to a single jet formed with the three partons, see the left panel in Fig.(41). Now you can think that the virtual corrections, with the three partons in the final state, will again fall into the same jet bin of the Born level process. When computing the real part, however, one has to consider the possibility of a collinear splitting that transforms parton 1 into two partons with lower energies. In this case, the algorithm will form two jets as shown in the right part of the Fig.(41). This is a phenomenon that a "good" jet algorithm should avoid in order to allow for theory-experiment comparisons beyond the leading order accuracy. One can modify the algorithm as follows: 1. all particles above a certain threshold are seeds, 2. combine all the particles in a cone to define a new proto-jet axis with

$$\phi_c = \frac{\sum_{i\in c} E_{T_i}\phi_i}{\sum_{j\in c} E_{T_j}} \qquad \eta_c = \frac{\sum_{i\in c} E_{T_i}\eta_i}{\sum_{j\in c} E_{T_j}} \tag{124}$$

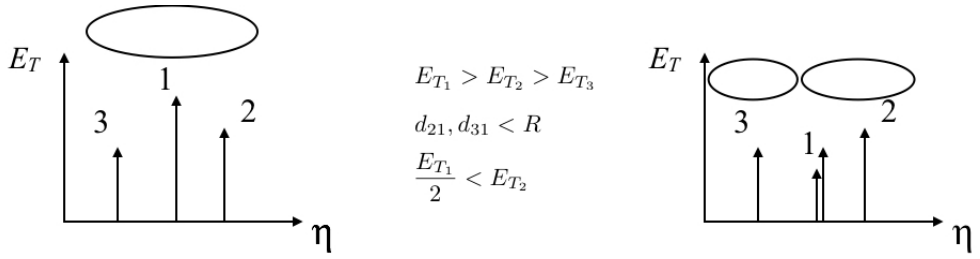


Fig. 41: A three partons configuration which is identified as a mono-jet event according to jet highest-E_T-seed algorithm (left panel). Collinear real emission from the parton 1 may lead to the production of two partons with lower energies such that, now, the algorithm identifies two distinct jets (right panel). The two collinear particles are separated in the two jets, spoiling the cancellation of the collinear singularity among the real and the virtual contribution which produces a mono-jet configuration as the leading order one.

and check the particles in the new cone: if they are the same, then a proto-jet has been formed (stable cone), 3. if proto-jets have small (large) overlap split (merge) them in some way introducing some technical parameters. In this way, jets number is now Collinear safe but unfortunately it is an InfraRed unsafe observable, that is a soft emissions change the number of jets as can be argued inspecting Fig.(42). An alternative to this problem could be the insertion of extra seeds at the mid-point of the stable cones, but

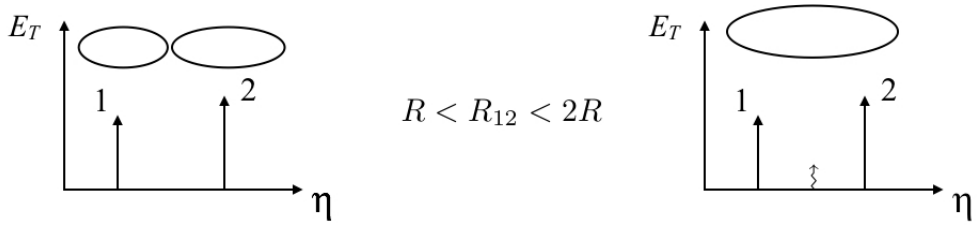


Fig. 42: First variant of the highest-E_T-seed algorithm described in the main text: the jets number is Collinear but non Infrared safe. Consider a leading order two-jets configuration as given in the left panel. Then, a soft real emission which occurs around the mid-point between the original two partons leads to the identification of a single jet (right panel). Indeed, there is now a large overlap between the two stable cones and they are merged into a single jet.

this is not a solution that turns out to be valid for higher order computations. A jet identification algorithm based on a cone definition that is fully InfraRed and Collinear safe is the seedless cone algorithm that works as follows: 1. consider a subset of particles, 2. combine momenta and search for a stable cone, 3. make it for all the subsets, 4. use split-merge as before. Nevertheless, it might become quite unpractical for very high multiplicity. Furthermore, it could happen that single particles with sufficiently high energy are not assigned to any jet (dark towers). Another solution is represented by the clustering algorithms that implement a sequential recombination along the line of the algorithms discussed for e^+e^- annihilation into hadrons. For the case of hadron collisions, the clustering algorithm requires the definition of two distances, that among each pair of proto-jets, d_{ij}, and the distance of each of them from the beams, d_{iB}. The algorithm works as follows: 1. for each pair compute d_{ij} and for each proto-jet compute d_{iB}, 2. if the smaller distance is one of the d_{ij} combine their momenta to form a new proto-jet ($p_{ij} = p_i + p_j$) while, if the smalled is one of the d_{iB}, the proto-jet i is promoted to a jet, 3. restart from point 1. This kind of algorithm is fully exhaustive and unambiguous (no need of split-merge procedures), it's also fast and does not generate dark towers. Introducing the quantities (caloremeter distances):

$$R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2 \tag{125}$$

the distances are defined in terms of an exponent p and another real parameter R, that acts as the radius of the cone algorithm, as:

$$d_{iB} = (q_{T_i})^p \qquad\qquad d_{ij} = \text{Min}\left\{(q_{T_i})^p, (q_{T_i})^p\right\} \frac{R_{ij}^2}{R^2}. \tag{126}$$

The choice of the value for the parameter p has an impact on the jet contours and the sensitivity to the detailed structure of the soft component of the event. For these reasons the preferred value (the standard) at the LHC experiments is $p = -1$ that define the anti-kt algorithm [23], other possible choices are the original kt algorithm [24] with $p = 1$ and the Cambridge/Aachen algorithm [25] with $p = 0$. The relevance of the second distance introduced, the one from the beams, appears when we analyse InfraRed and Collinear safety of the jet definition. Collienar safety is guaranteed by the fact that $R_{ij} \to 0$ in the collinear limit, while for InfraRed safety the point is that soft real radiation is either assigned to a jet or forms a jet by itself. In the latter case, the "soft" jet, being below any measurable threshold, is combined with the virtual counterpart, without spoiling the cancellation between real and virtual contributions. We conclude this section by observing, once again also for hadron collisions, that even if we are able to build InfraRed and Collinear safe algorithms for jet identification, the number of jets we identify and how they are distributed depends upon how we search for them, i.e. the algorithm we use (and its parameters). Anyway, this does not prevent us to test the perturbative regime of QCD probing the theory in high energy scattering processes among the elementary constituents.

3.4 Jets phenomenology

Let's start by considering the inclusive jet cross section at the LHC. In perturbation theory it is computed through the formula:

$$\frac{d^2\sigma}{dydp_\perp^2} = \sum_{i,j,k} \int dx_1 dx_2 f_i(x_1,\mu) f_j(x_2,\mu) \frac{d^2\hat{\sigma}_{ij\to kX}}{dydp_\perp^2}. \tag{127}$$

While integrating each matrix element (and each subtraction counterterm needed order by order in perturbation theory) over the proper phase space, the event kinematics is passed through a jet reconstruction algorithm, i.e. there is a convolution with one of the jet finder functions described in the previous section, then the experimental constraints of the measurements (cuts) are applied. The error associated to the perturbative part of the prediction can be estimated from the variations of the unphysical scales plus the error related to pdf's and strong coupling determination. One then can estimate non perturbative corrections using Monte Carlo generators, first evaluating the ratio (bin-by-bin) of the Monte Carlo cross sections with and without hadronisation, and then multiplying by this ratio the NLO parton-level cross sections. The uncertainty related to this procedure is estimated as the maximum spread of the correction factors obtained using different Monte Carlo programs. In Fig.(43), the excellent agreement is shown among theory [26] and experiment [27] for the doubly differential distribution of the inclusive jet cross section. Left and right panels refer to different choices for the R parameter characterizing the size of the jet in the anti-kt algorithm, and also within this variation the agreement is excellent over many orders of magnitude! Recently, the Next-to-Next to leading order computation for two jet production at hadron colliders has been performed in the leading color approximation [28] that amounts to selectively compute all the terms that are enhanced by colour factors and for this reason are supposed to dominate. In Fig.(44) it is shown the excellent agreement among theory and experiment over almost seven orders of magnitude for the invariant mass distribution of the dijet pair in different rapidity bins, at the LHC.

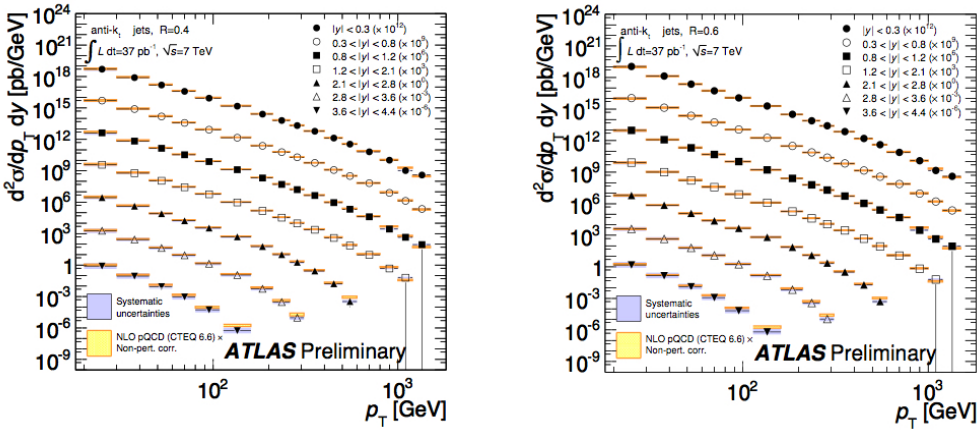


Fig. 43: Data-theory comparison of the double-differential inclusive jets cross section (in p_T, y) by the ATLAS collaboration. The jets are reconstructed according to the anti-kt algorithm setting $R = 0.4$ (left panel) and $R = 0.6$ (right panel).

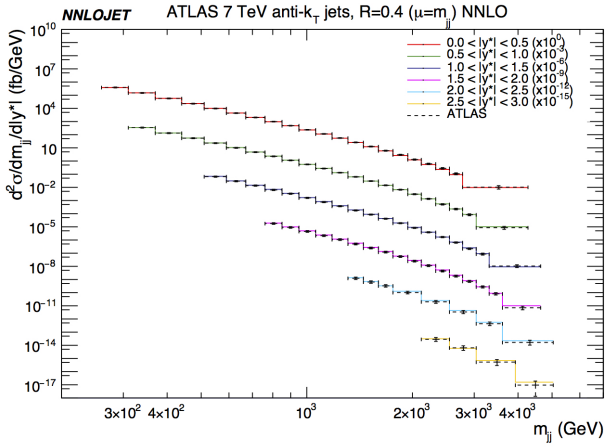


Fig. 44: Comparison between data and the state of the art NNLO prediction (in the leading colour approximation, using the NNLOJET code) for the double-differential inclusive jet cross section (in m_{jj}, y) by the ATLAS collaboration. The agreement is excellent over seven orders of magnitude.

4 QCD at higher and all orders

The purpose of this section is the introduction to those methods to treat QCD radiation in the perturbative regime that allow to reach the highest accuracy in testing this theory at colliders. Very few technical details are provided. The emphasis is more on the ideas underlying these methods, with the aim to give an orientation on what can be expected from their use and what cannot.

4.1 Higher order corrections for LHC processes

The exact computation of a hard scattering cross section including all the contributions up to a given order in the coupling constant goes under the name of fixed order computation. Although the problem has been solved at NLO, the complications that show up going at higher orders is formidable and we are still quite far from automation at NNLO. This is the frontier from the computational point of view. These computations are revealing new mathematical structures and a big effort is currently required also to the

mathematicians, so that the situation is evolving fast. Note also that, in our knowledge, there is no higher order computation performed so far that has not become relevant for a comparison among theory and experiment. Indeed, we can say that every time a new higher order computation has been performed a new phenomenological challenge for the Standard Model is ready. Recently, an N3LO QCD calculation has been also completed. It is the cross section for the Higgs boson production in gluon fusion at the LHC [29, 30], see Fig.(45) for representative diagrams up to the NLO. In the Standard Model there is

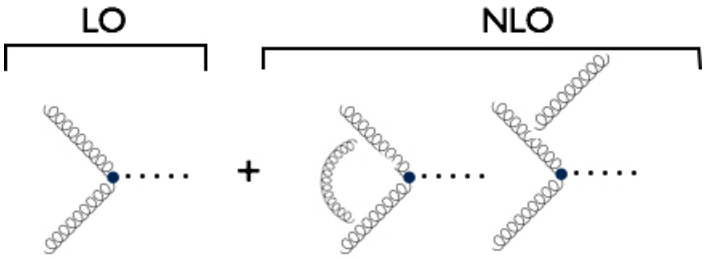


Fig. 45: Selection of diagrams contributing to the production of a Higgs boson in the gluon fusion channel up to NLO in the strong coupling. The interaction between the gluons and the Higgs is given by the effective vertex obtained pinching the dominant one-loop diagram given by the top quark triangle (large m_t approximation).

not a vertex coupling directly the Higgs boson to gluons. This coupling in the SM proceeds, with very good approximation, via a loop of top quarks. From such top loop diagrams, where both gluons and a Higgs boson are attached, it is possible to deduce an effective theory including the Hgg vertex shown in Fig.(45). Going to NNLO, one has to include three new cathegories of diagrams: the ones with two radiated partons (double real), the ones with one radiated and one virtual (real-virtual one loop diagrams) and the ones with two virtual diagrams (double virtual or two loops). Needless to say that going to N3LO the categories increase and they range from three radiated real partons up to three virtual partons. For the time being this computation in [29, 30] has been performed at the inclusive level only, that is only the total cross section has been predicted at N3LO. The experiments at CERN on the other hand measure the cross section (even differentially) but only in a certain fiducial volume. Therefore, the comparison among theory and experiment requires an extrapolation of the measured result over the unexplored region, that is done with Monte Carlo tools. In Fig.(46) [29], the result of the N3LO computation as a function of the common choices for the renormalization and factorization scales is shown. Here one can appreciate the flattening of the scale dependence going from one order to the next. We usually say that scale variations provide the uncertainty from the lack of knowledge of missing higher orders. Nevertheless, there is not a unique prescription on how to do it. This is a point debated since long. Given that these scales are fictitious and the full theory does not depend on them, what is obvious is that the scale dependence of a prediction at a certain order must exactly match the scale dependence of the collection of all the higher order contributions up to infinity with the opposite sign. For this reason, one can reasonably think that scale variations provide a taste of the size of the missing higher orders. At the differential level, NNLO computations in hadronic collisions are available for some time for the Higgs boson and single gauge boson production and more recently also for every pair of bosons and for top anti-top production. These computations allow to make very precise tests of the Standard Model at the LHC and at the same time, of course, set the scene to disentangle any possible new physics signal at this machine.

4.2 Resummation

At the differential level, fixed-order predictions might be affected by logarithmic enhancements which show up when a particular restricted region of the space space is considered. Such logarithmic enhancements can make meaningless the fixed order prediction in that phase space region so that apparently we assist to a break down of the perturbation theory. For a reliable comparison with experiments, whenever

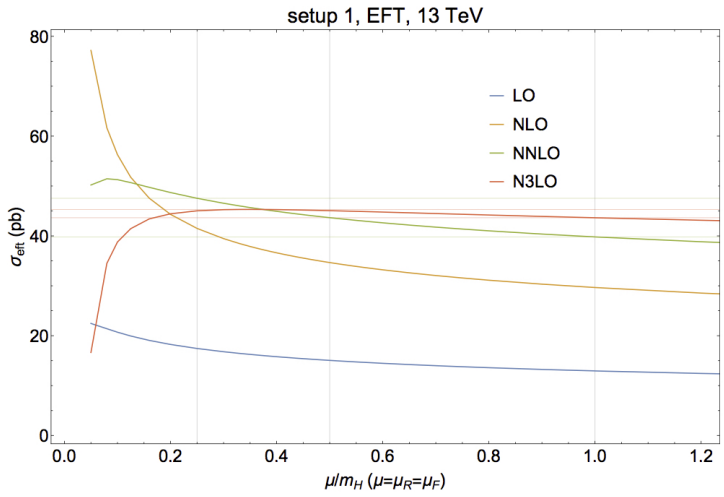


Fig. 46: Comparison of predictions with increasing accuracy up to N3LO for the Higgs production cross section in the gluon fusion channel as a function of the scale $\mu = \mu_F = \mu_R$. The inclusion of higher order corrections not only improves the estimate of the total rate but nicely reduces the scale variation. At N3LO, the dependence is almost flat in a wide range of scales, giving a solid confirmation of the reliability of this result.

possible, a consistent procedure of "resummation" of the logarithmic enhancements to all orders is required. We have already encountered the UV log and the Renormalisation Group Equation that allows to effectively resum this log producing the observed running coupling. Indeed, the resummed expression for the running coupling has to be used in order to connect, for example, the phenomenon of hadronic decay of the τ lepton and the hadronic decay of the Z boson. Starting from $\alpha_s(m_\tau) \sim 0.36$ the running nicely predicts:

$$\alpha_s(M_Z) = \frac{\alpha_s(m_\tau)}{1 + b_0\,\alpha_s(m_\tau)\log\frac{m_Z^2}{m_\tau^2}} \sim 0.12 \tag{128}$$

which is in good agreement with measurements. The reduction of α_s by factor of 3 implies, roughly speaking, that:

$$b_0\,\alpha_s(m_\tau)\log\frac{m_Z^2}{m_\tau^2} \sim 2 \tag{129}$$

and proves that:

$$\alpha_s(M_Z) \neq \alpha_s(M_\tau)\left(1 - b_0\,\alpha_s(m_\tau)\log\frac{m_Z^2}{m_\tau^2}\right) < 0. \tag{130}$$

The expansion of the running coupling to the first order, given by the expression in rhs in the above formula, is even negative, a completely non sense result! So that the resummation of the log is mandatory. Then we have encoutered the collinear log in collisions with an hadron in the initial state that is effectively resummed by the Altarelli-Parisi equations. The common feature in the above examples is that they are both multiscale problems, with two (or more) scales, far apart each other. In the first case, they are given by m_τ and M_Z, while in the second case the scale of the hard process Q and the scale of the strong coupling regime of QCD. In these situations, multiple emission processes connecting the two scales spoil the convergence of the perturbative expansion and cannot be neglected. From another perspective, large logarithmic contributions can remain in the computation, despite the fact that more inclusive quantities (as the total accepted signal) are still finite, if the experimental cuts produce an unbalanced cancellation of real and virtual contributions, with a suppression or emphasis of the real contribution. Think for example once again to the jet cross section in e^+e^- annihilation into hadrons at NLO. Inclusively, the

real and the virtual contributions together give a finite result, but if we restrict too much the value of the jet resolution parameter y_{cut}, the three jet cross section fraction increases more and more becoming grater then 1, while at the same time the two jet fraction becomes negative! This situation is of course untenable and it signals the unbalance we meant above. Fixed order perturbation theory has to be used only far from these kinematical regions. But, what to do if experimentally one can perform good measurements also assuming such small values of the resolution parameter? This problem, on the theory side, can happen for every observable which is sensitive to soft and/or collinear radiation and the situation does not improve that much including only a finite number of higher orders contributions. Consider for example the computation of the transverse momentum distribution of a vector boson produced in hadron-hadron collisions. Illustrative diagrams of this process are shown in Fig.(47).
The scale of the process is of course the mass of the vector boson (M). The schematic behaviour of the perturbative expansion for the p_T distribution of the boson is shown in Fig.(48). At leading order (α_S^0), this distribution is just a δ function selecting zero p_T. Experimentally the measurement will have a certain resolution and the result will be given as an histogram as for the theory prediction. At Next-to-Leading order (α_S), the logarithm (L) of the ratio M^2/p_T^2 shows up with a positive diverging behaviour in the low p_T limit. Note as well the presence of a characteristic feature of the fixed-order computation. Since the area under the histogram represents the total cross section (which is a Collinear and InfraRed safe observable), the first bin is negative and if we restrict the bin it tends to become divergent to compensate the divergent positive behaviour of the integral of the rest of the distribution. At the next order (α_S^2) the distribution tend to minus infinity approaching low p_T and up to two more powers of the same $\log$ are present. Once again, the first bin is large and positive and compensates to produce a finite total rate. One can show that at order n all powers m of the $\log$ with $0 < m < 2n-1$ manifest. This situation repeats also for other measurements that are very well defined experimentally like the small mass limit of a jet in e^+e^- annihilation, the small jet radius in hadron collisions or the large transverse momentum distribution of massive quarks produced at high energy and so on. Predictions for all these observables need resummation of large logs (L) in the kinematical region where indeed $\alpha_s L^2 \sim 1$. The contributions have a well defined structure and can be schematically organized as shown in Fig.(49).
We do not resum the whole series, just classes of contributions. The sum of all the terms in the first column in Fig.(49) is called the Leading Log (LL) resummation, the second column the Next-to-Leading Log (NLL) resummation and so on. This time the next term in the series, the next column, will provide a relatively lower contribution because it will be suppressed by one power of L with respect to the previous one. Note that resummed results can be combined with the fixed order "horizontal" contributions via a so called matching procedure that removes double counting of the logarithmic contributions. However, whether the series converges or not (and if it has an exponential form) depends on the quantity being measured. For examples, the $\log(y_{cut})$ contribution in e^+e^- annihilation does not exponentiate for the JADE algorithm but it does for the kT algorithm! In general, for the computation of the coefficients of the series, it is not needed to compute the whole correction at a given order, it is enough to select and compute those contributions that catch the soft and collinear behaviour up to the desired order. Furthermore, it is not needed to compute explicitly all the coefficient in a certain column, from a subset of coefficients it is possible to deduce the all order expansion. Resummation for basic variables measurable at the LHC, such as for example the transverse momentum of the Higgs boson and vector gauge bosons, have been

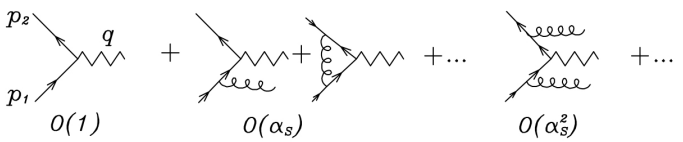


Fig. 47: Illustrative diagrams contributing to the production of a vector boson in hadron-hadron collisions.

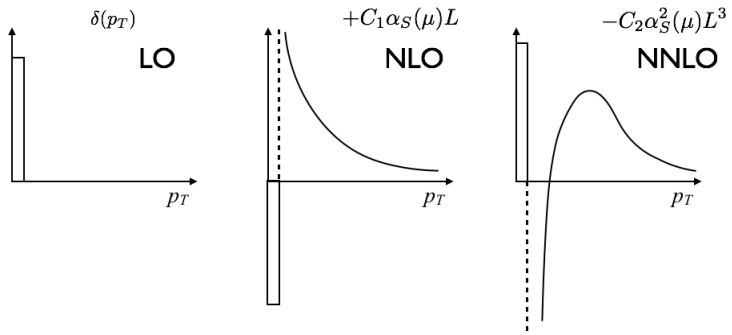


Fig. 48: Fixed-order predictions including an increasing content of radiative corrections for the differential p_T distribution of the vector boson. A logarithm $L = \log M^2/p_T^2$ shows up at NLO, its square at NNLO and so on. The predicted behavior in the small-p_T limit is divergent and so completely unphysical.

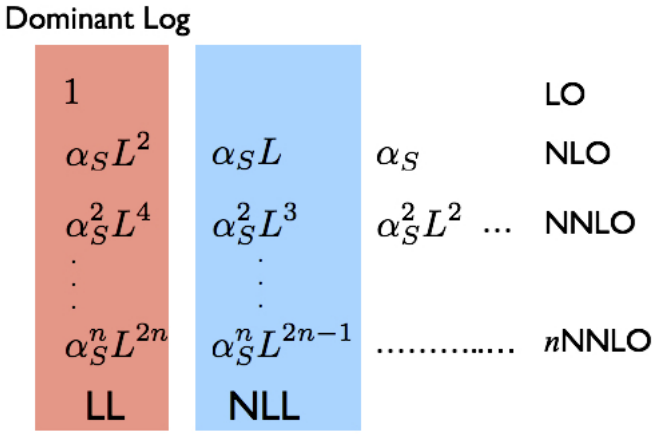


Fig. 49: Structure of the logarithms which occurs at each order in perturbation theory.

performed recently including up to N^{3}LL+NNLO corrections [31]. The result for the Higgs boson is shown in Fig.(50).
While the fixed order results have no physical meaning in the small p_T region, the resummed one has an acceptable behaviour in agreement with the experimental results (not shown in the figure) that present a maximum and then goes down at lower p_T. Note also the strong reduction of the scale variation that is a typical benefit of the procedure for that part of distributions that are pathological for the fixed order computation.

4.3 Parton shower

As we have seen in the previous section, the resummation of higher order logarithmic enhancements contributing to single variable distributions has reached an high level of accuracy for the cases studied. There is another possible strategy which allows to resum only the leading soft and collinear logarithmic enhancements but, on the other hand, is fully differential. This means that it is not limited to the single inclusive distribution and can be used for all the observables for a given process. Starting from a hard 2 to 2 process, it is possible to build an approximation for all the diagrams in which subsequent multiple emissions have distributed the available energy to a shower of partons. To give an idea about how this works we will consider again e^+e^- annihilation into hadrons. In QCD we aim at a description as given in Fig.(51).
We start from the $q\bar{q}$ pair produced at high energy. It happens at a certain fixed starting hard scale Q.

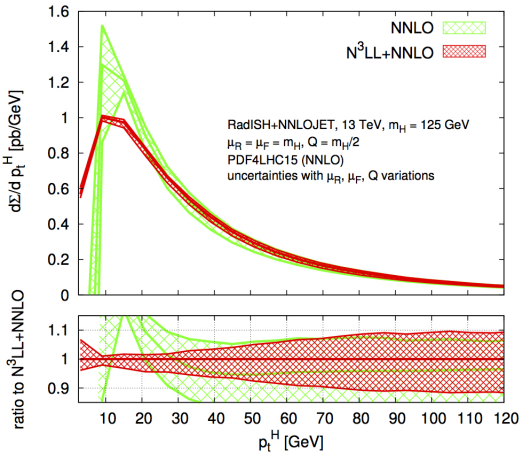


Fig. 50: Comparison between the NNLO fixed-order prediction and the N^{3}LL+NNLO one for the differential p_T distribution of the Higgs boson in gluon fusion channel. As expected, at high p_T there is an excellent agreement, while in the small p_T region, the resummed prediction gives a reliable and physical result, while the fixed order prediction is divergent. Moreover, resummation nicely reduces the scale variation, especially in the small p_T limit.

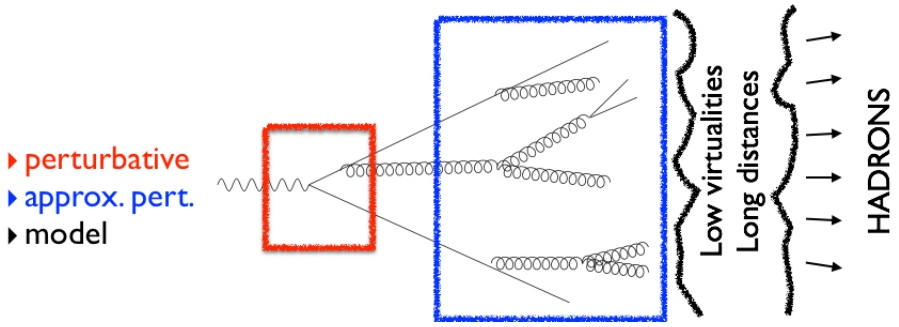


Fig. 51: Sketch of the QCD evolution from the hard scattering to the final state hadrons in the simple case of e^+e^- collisions. The main behaviour of the shower of partons can be captured by an approximate perturbative QCD description, which is implemented in the so-called Parton Shower programs.

Then we look for an iterative formula to approximately describe the sequence of parton splittings starting from the real contribution. What we are going to build is an effective description of the shower process in terms of splitting functions and propagators and not the full matrix element squared for every possible history, which is a task still beyond the possibilities of any available computer. Nevertheless, it can be shown that such construction catches the main features of the QCD evolution of the external partons that took part to a given hard process, from the hard scale of the process down to scales a bit above 1 GeV. To start with, we know that collinear emission from each leg is enhanced and we can use Eq.(47) and Eq.(48) to describe the emission probability as:

$$d\sigma_R \sim \sigma_0 \sum_i \frac{d\theta^2}{\theta^2}\, dx\, P_{ij}(x). \tag{131}$$

Of course, this probability still diverges. We observe that to describe the collinear limit we could have used also other variables that are good as well. Indeed, we get the same limiting behaviour if the new variable is proportional to θ^2. Consider for example the invariant mass of the internal propagator q^2 in a splitting, whose expression in the collinear limit is $q^2 = x(1-x)\theta^2 E^2$, or the transverse momentum associated to the emission, $k_\perp^2 = x^2(1-x)^2\theta^2 E^2$. We would get identical results for the collinear limit,

but different extrapolations away from it. Let us consider the virtuality of the internal line so that we have:

$$d\sigma_R \sim \sigma_0 \sum_i \frac{dq^2}{q^2}\, dx\, P_{ij}(x). \tag{132}$$

Still, of course, the probability for a given emission path, which is obtained by reiterating (i.e. exponentiating) this formula, can well exceed 1! While we want that the sum of the probabilities of all shower configurations must yield 1. Indeed, approaching the strict collinear limit, the fixed order theory prediction is challenged and the enhancement, phenomenologically, are not really there as such. In turn, we need to damp this emission probability with a factor that has to restore unitarity keeping into account the missing virtual corrections and, at the same time, resumming the logarithmic enhancement. To address this point, while discussing the basic idea, we make a further simplification by assuming that we have just one kind of parton (the gluon for example) and so one kind of splitting P_{gg}. At some point of the shower the number of gluons will certainly prevail on the number of quarks, nevertheless our simplifying assumption will obviously not be a good approximation for quark initiated processes, and it has to be considered just as an illustrative example. The reader interested in the details of the most general case can go through excellent reviews of the subject [32, 33]. To restore unitarity we normalise the emission probability involving an internal propagator with virtuality q^2 given in Eq.(132) multiplying by the inclusive probability of no emission of radiation implying and internal propagator virtuality from the maximum available virtuality Q^2 down to the virtuality q^2:

$$d\sigma_{first\ emission} = P_{gg}(z)dz\frac{dq^2}{q^2}\Delta_g(Q^2, q^2). \tag{133}$$

In the formula above, we do not include the leading order cross section σ_0 because we are now considering emission from a single final state parton (gluon in our simplified exercise), so that we are building the iterative factor that describes the emission probability from each leg in a Markov chain like process. The factor Δ_g is called a Sudakov form factor. Of course, this "no emission" probability has to be related to the emission probability, it has to be a function of it. Given that for smaller and smaller q^2 the emission probability diverges, in this limit the inclusive "no emission" probability has to go to zero rapidly, so regularising the divergence. Now, let us compute Δ (here comes the resummation!). In the dominant collinear limit, the probability of emission implying an internal propagator virtuality among q^2 and $q^2 + dq^2$ is given by:

$$dW_g = \frac{dq^2}{q^2}\int_{z_{min}}^{z_{max}} dz\, P_{gg}(z) \tag{134}$$

so that, for such an infinitesimal range of virtuality dq^2 the "no emission" probability is given by the unitary condition:

$$dP(\text{no emission}) = 1 - dW_g = 1 - \frac{dq^2}{q^2}\int_{z_{min}}^{z_{max}} dz\, P_{gg}(z). \tag{135}$$

Now, to get the (resummed) "no emission" probability over the finite range from q^2 up to Q^2, we do not integrate this expression, but instead we take the product over "all" the infinitesimal paths, which we consider as independent events. In doing so, the probability of "no multiple emission" of every number of gluons turns out to be:

$$\Delta_g(Q^2, q^2) = \prod dP(\text{no emission}) = \exp\left\{-\int_{q^2}^{Q^2}\frac{dk^2}{k^2}\int_{z_{min}}^{z_{max}} dx\, P_{gg}(z)\right\}. \tag{136}$$

Note that in the limit of small virtualities we know that all order emissions are relevant and that factorization holds. This product has an exponential form and produces the damping behaviour we were looking for. Indeed, it damps the emission probability at low virtualities and it does not count much at

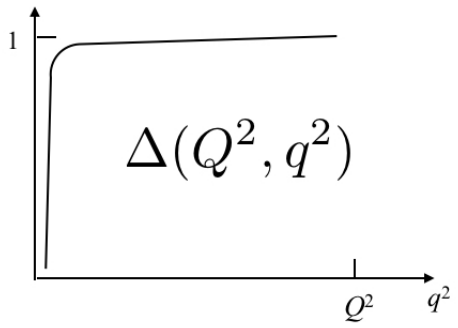


Fig. 52: Characteristic behaviour of the Sudakov form factor or "no emission" probability. It rapidly vanishes in the $q \to 0$ limit, while saturates to 1 at high virtualities.

large virtuality. A typical form is shown in Fig.(52). In our simplified QCD model, the parton shower is then produced iteratively generating radiation according to the probability distribution:

$$d\sigma_{first\ emission} = P_{ji}(z)dz\frac{dq^2}{q^2}\exp\left\{-\int_{q^2}^{Q^2}\frac{dk^2}{k^2}\int_{z_{min}}^{z_{max}}dzP_{ji}(z)\right\}. \tag{137}$$

At each iteration of this formula one updates the upper Q^2 scale to the extracted q^2 of the previous step, until the virtuality of the internal line reaches the cut-off scale. We stress that by construction unitarity is restored, so that the event simulation will not alter the value of the original cross section at the starting point (σ_0 at the origin). A number of observations have to be done. In real life event generators, the qualitative reasoning to describe the iterative procedure implemented in a parton shower given above are supplemented by a large number of technicalities we have not discussed. They go from the implementation of the momentum reshuffling after each emission to the management of multiple splitting processes. Another complication is related to the treatment of hadrons in the initial state. In that case one has to take into account the structure of the colliding hadrons, but the shower is worked out with the same logic (in that case one speaks of backword evolution and space-like shower). Furthermore, for a parton shower implementation there is also a number of choices to do, like: the evolution variable (that we have set to the the virtuality of the internal propagator), the scale to be used in the strong coupling, the cut-off scale. As for the evolution variable, it has a relevant impact. Indeed, note that the iterative multiple emission description obtained starting from the collinear approximation discussed above cannot represent adequately multiple soft emissions. It is true that damping the virtuality of the internal propagator or the tranverse momemtum of the emission also the soft radiation goes out, but we also know that the divergent behaviour of soft radiation has a universal different factorized form. From direct computation of QCD amplitudes, it turns out that for soft radiation there is a coherent effect such that, if one chooses as evolution variable the angle or the transverse momentum, the procedure automatically catches also the leading logarithmic enhancement associated to the soft radiation. This is a very good news because, if also soft radiation has to be more and more collinear, we can imagine that, although in principle it has a long range, in QCD it stays relatively close, so that the hadronization can be considered and modelled as a local process. An important application of a parton shower at the LHC is the study of the jet shape. Such a study requires of course the resummation of large soft and collinear logarithmic enhancements. A relevant variable in this case is the fraction of the jet transverse momentum contained within a ring of tickness Δr around the jet core, $\rho(r)$ (averaged over an ensamble of jets), with:

$$r_i = \sqrt{(\Delta_{i,jet}y)^2 + (\Delta_{i,jet}\phi)^2} \tag{138}$$

$$r_a = r - \Delta r/2 \qquad r_b = r + \Delta r/2 \tag{139}$$

$$\rho \quad = \quad \frac{1}{N}\sum_{jets}\frac{1}{\Delta r}\frac{\sum_{r_a<r_i<r_b} p_{Ti}}{\sum_{r_i\leq R} p_{Ti}}. \tag{140}$$

In Fig.(53), we can see the general agreement among the measurement of $\rho(r)$ distribution and parton shower Monte Carlo programs. Another intersting variable is the fraction of the jet transverse momentum

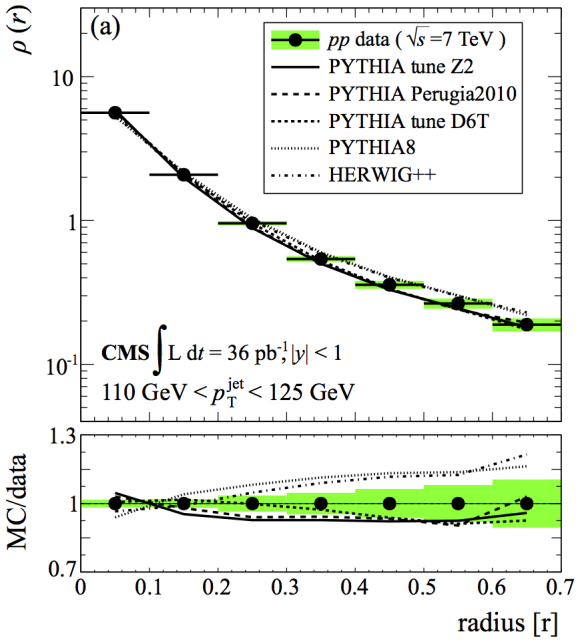


Fig. 53: MC/data comparison for the jet shape variable ρ representing the fraction of the jet transverse momentum in a ring of tickness $\Delta r = 0.1$, measured by the CMS collaboration.

contained within a circle of radius r around the jet core, $\Psi(r)$. In Fig.(54) the ATLAS measurement for $1-\Psi(0.3)$ is shown. In this case, we see that in principle this variable could be used to distinguish jets

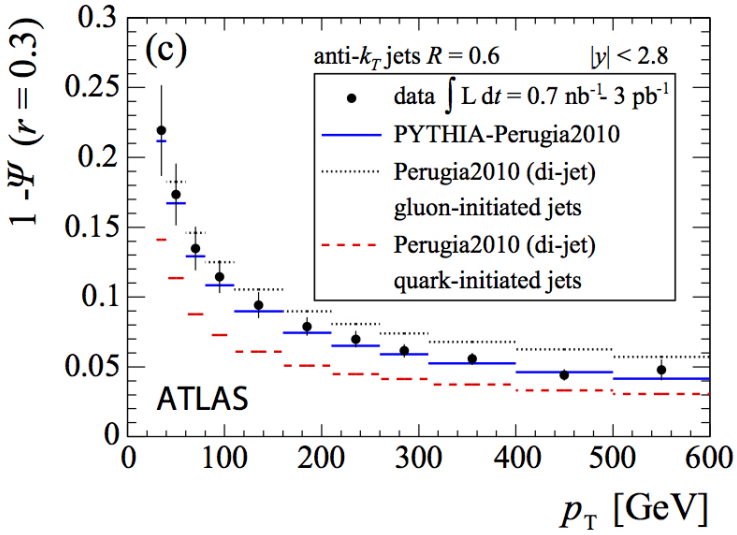


Fig. 54: MC/data comparison for the jet shape variable $1-\Psi(r)$ with $\Psi(r)$ representing the fraction of the jet transverse momentum in a circle of radius r (for the tipical value $r = 0.3$), measured by the ATLAS collaboration.

initiated by a gluon or a quark.

4.4 Merging, matching and both

Shower Monte Carlo programs effectively handle and resum the multiple collinear and soft QCD emissions and they provide the full event simulation, including the conversion from partons to hadrons. For this reason they are an essential tool to study the sensitivity of the experiments in high energy physics. On the other hand, they are able to generate cross sections for the requested hard process that are correct only at leading order. Furthermore, the soft and collinear approximations underlying the parton shower (PS) evolution may fail to reproduce the full pattern of hard wide-angle emissions at higher orders. One would like to include in a parton shower simulation all relevant hard processes while taking advantage of available higher order corrections. We are indeed on this path. The first step forward has been the merging of multi-jet final states preserving LO+LL accuracy, done combining together all the LO matrix elements with up to n partons in the final state. The problem is how to regularize the divergent cross sections in the soft and collinear limits of the corresponding matrix elements (for $x > 2$, x being the number of partons in the finale state) and to generate hard multijet configurations in a way that is not very sensitive to the regularization prescription. Here we will not give any proofs, but we will just outline the basic idea about how it works in practice. To this aim, let us consider again, for the sake of simplicity, the case of e^+e^- annihilation at a cm energy Q. There have been more then one proposal on how to reach this goal [34, 35], here we limit to show how the Catani, Krauss, Kuhn and Webber (CKKW) algorithm [34] works. The first step is the introduction of a merging scale, so: 1. choose a scale Q_1 which fixes the kT algorithm resolution parameter $y_{cut} = Q_1^2/Q^2$. This parameter acts as a cut-off regulator for the computation of the m-jet cross section using the exact tree level matrix element with m partons in the final state. This means that any m-partons configuration that would be clustered into a number of jets lower than m is given zero weight. Repeat the procedure for every m from 2 to n (typically $n = 5$ or 6), so that for each multiplicity m and partonic subprocess i (for each multiplicity there is finite number of them increasing with multiplicity) you now have a corresponding finite numerical value of the cross section $\sigma^{(0)}_{m,i}$, where the (0) apex refers to the fact that we are considering leading order cross sections. 2. Select a jet multiplicity and a partonic subprocess with probability:

$$P(m,i) = \frac{\sigma^{(0)}_{m,i}}{\sum_{kj}^{k=n} \sigma^{(0)}_{k,j}}. \tag{141}$$

3. Generate an event with probability given by the corresponding exact matrix element. 4. Reconstruct the PS probability of this event first using the kT algorithm backward, recombining partons until only 2 remain, and then building an event weight combining Sudakov form factors (as propagators) and splitting probabilities (as vertices) with the strong coupling evaluated at the scale of the branching process for each branch. 5. Accept or reject the event according to the combined PS weight. If the event is accepted, assign it a color configuration and start the shower from each leg of the m partons event. The initial scale of the shower should be set to the scale value of the node at which that parton was created. Furthermore, during the shower one has to veto radiation with transverse momentum larger then Q_1. If the event is rejected go back to point 3. Using the steps listed above, it has been proved that the results do not depend on the cutoff scale Q_1. Of course, the algorithm needs modifications when there are hadrons in the initial state, but the logic is the same: the idea is again to regularize the matrix elements separating the phase space into hard (ME domain) and soft (PS domain) and then to reject or accept an event, and eventually start the shower, according to the PS weight of the event. In Fig.(55) and (56), we can appreciate the adequate description given by the shower Monte Carlo programs for both inclusive and differential multijet cross sections at the LHC. However, one has not to expect too much from such a PS description. Note, for example, that a 5% uncertainty on the strong coupling induces a 35% uncertainty on the 6 jets production rate, on top of which one has to consider scale variations (both renormalization

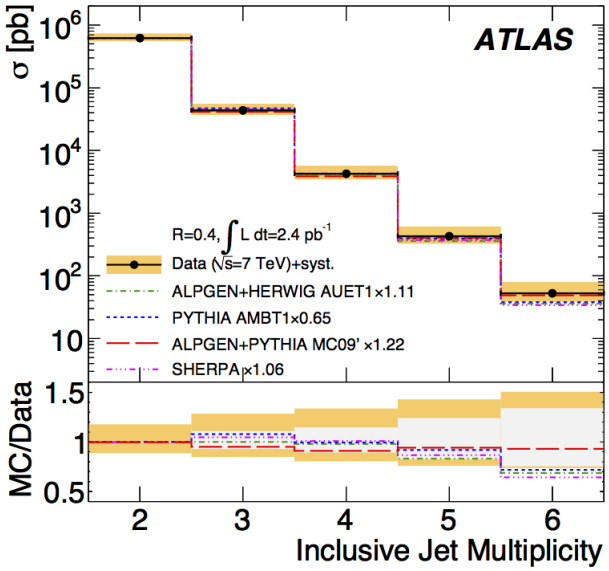


Fig. 55: Measurements of the total cross section for different inclusive jet multiplicities by the ATLAS collaboration compared to different MC generators.

and factorisation) and hadronization effects. So, in turn, the description of multijets final states with merging shower MC programs might be considered more as an exercise to tune those tools than a test for perturbative QCD.

The matching of NLO computation with PS LL resummation has been a big progress of the recent years. The first proposals have been the MC@NLO [36] and POWHEG [37] methods. Let us take for example the POWHEG method. The acronym stands for positive weight hardest emission generator. The technical problem to solve was how to avoid the double counting of radiation described by the NLO computation and the shower process. In particular, we want to keep the description of the hard and wide angle radiation as given by the NLO computation and exploit the resummation performed by the shower Monte Carlo program for the soft and collinear radiation. In this way, the full differential description offered by the PS programs will gain the normalization and a formal accuracy, for sufficiently inclusive observables, driven by the NLO calculation. The interested reader can look into the original reference for the proofs [37], here we just give a sketch of the construction in practice. We start considering the first emission in a parton shower approach:

$$d\sigma_{\text{first emission}} = \sigma_0\, dz \frac{dk_\perp^2}{k_\perp^2} P_{ij}(z)\Delta_j(Q^2, q^2). \tag{142}$$

To have NLO accuracy on inclusive variables, we start generating the "underlying" Born kinematic configuration with probability given by the parton level NLO computation. So we promote σ_0 to:

$$\bar{B}(\Phi_B) = B(\Phi_B) + V(\Phi_B) + \int d\Phi_{rad}[R(\Phi_B, \Phi_{rad}) - C(\Phi_B, \Phi_{rad})] \tag{143}$$

In this formula, B and R are the Born and the Real matrix element squared, C is the counter-term that makes finite the Real cross section and V is the combination of the Virtual matrix element and the integrated counterterm, along the lines of the discussion in the first section. Actually, one should partition the Real matrix element into pieces that are each singular in just one kinematical radiation region and build as many of the last piece in the formula above as singular regions. Their simultaneous treatment is addressed with Monte Carlo techniques and it does not pose a problem. For ease of notation,

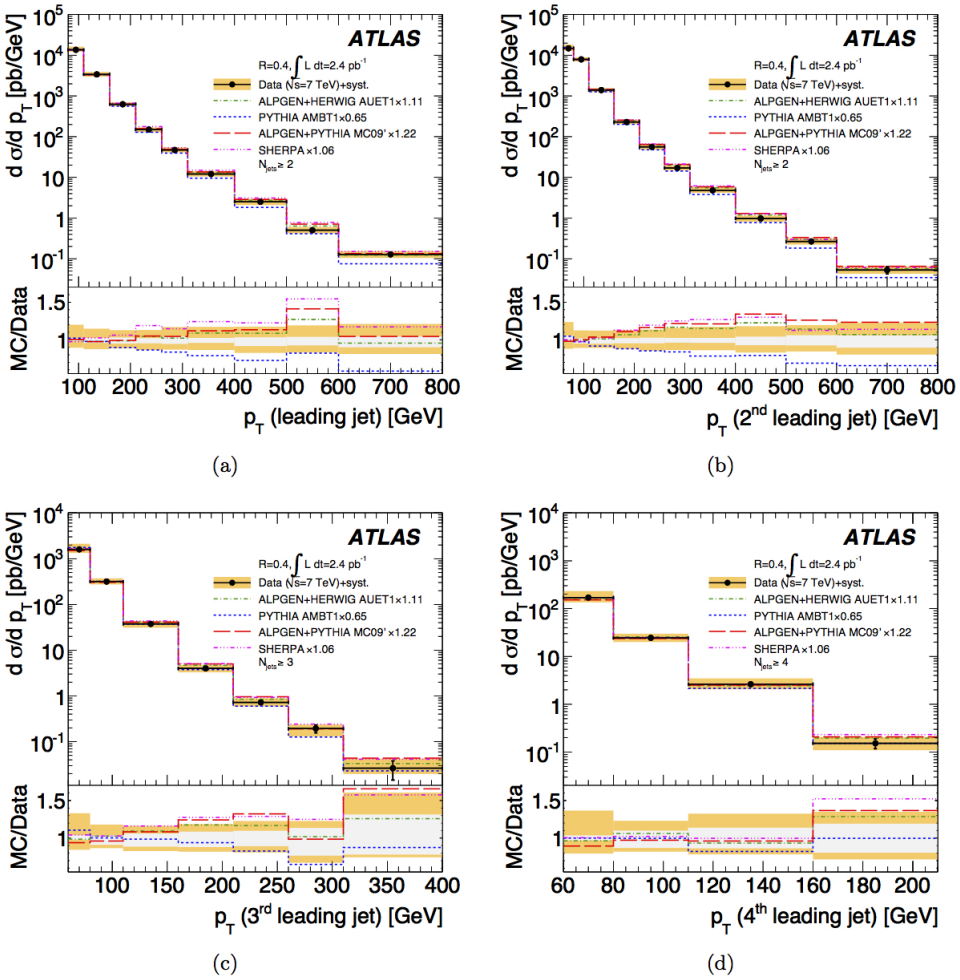


Fig. 56: Measurements of the p_T distribution of the leading jet (a), of 2^{nd} leading jet (b), of 3^{rd} leading jet (c), of 4^{th} leading jet (d) by the ATLAS collaboration compared to different MC generators.

here we leave just one of them. Furthermore, with Φ we indicate the phase space variables for a Born configuration (Φ_B), a Real one (Φ_B, Φ_{rad}) or just the radiation variables (Φ_{rad}). To match the Leading Logarithmic resummation performed by the parton shower, while at the same time avoiding double counting, one has to generate the first radiation using the formula

$$d\sigma = \bar{B}(\Phi_B) d\Phi_B \, d\Phi_{rad} \frac{R(\Phi_B, \Phi_{rad})}{B(\Phi_B)} \exp\left\{ \int \frac{R(\Phi_B, \Phi'_{rad})}{B(\Phi_B)} \theta(k'_T - k_T) d\Phi'_{rad} \right\}, \tag{144}$$

then assign a color configuration and shower the event vetoing emissions harder then the first one. We see that the damping Sudakov form factor is built upon the ratio of Real and Born matrix elements and not just by its divergent limit represented by the splitting functions. On the other hand the exponential resummation is certainly valid in such limits so that for the method to work properly one has to ensure that soft and collinear configurations are the only sources of enhancement for the ratio R/B. In the first method proposed for the NLO+PS matching (MC@NLO), the procedure is a bit different. In that case, the mapping used in the NLO calculation to connect a Real configuration to a Born one plus the radiation variables is chosen to be the one implemented in the Parton Shower program to describe the emission. Then the PS weight associated to the configuration of the real event (called Hard in the MC@NLO framework) is subtracted from the value computed with the exact Real matrix element and added back to the weight of the corresponding (mapped) Born configuration (called Soft). Note however that, being all the

contributions finite, it is always possible to built unweighted event samples before showering. The two procedures MC@NLO and POWHEG are then different in the way they treat subleading contributions and, in general, an estimate of the size of the uncertainty related to the matching procedure can be represented by the differences among the results found simulating the same process with the two tools. As an example of the results obtained at NLO+PS accuracy, we show the case of top anti-top production in proton proton collisions at 14 TeV [38]. As shown in Fig.(57), the MC@NLO generator reproduces very well the NLO distributions that are inclusive over extra QCD radiation as the rapidity and the transverse momentum of the top quark. In the case of variables that are sensitive to radiation attached to the Born

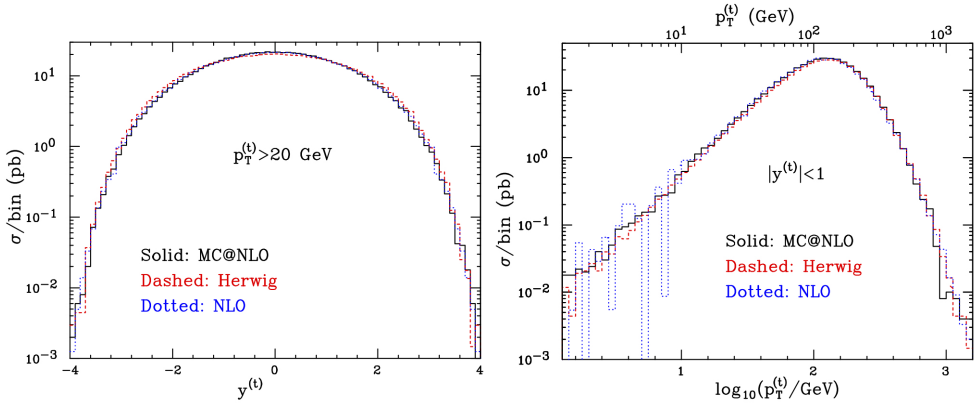


Fig. 57: The NLO+PS(MC@NLO) prediction for the single inclusive observables $y^{(t)}$ (left) and $p_T^{(t)}$ (right) (rapidity and transverse momentum of the top) in top-antitop production at LHC14 is compared with the PS(Herwig) and the NLO fixed-order predictions. Here, a very good agreement is found among all the prediction.

process, as the transverse momentum of the top anti-top pair in Fig.(58), we observe a complementary behaviour of the NLO and MC approaches regardless of the cuts on the rapidities and transverse momenta of the top quark. In the tail of the transverse momentum distribution of the top pairs, the NLO

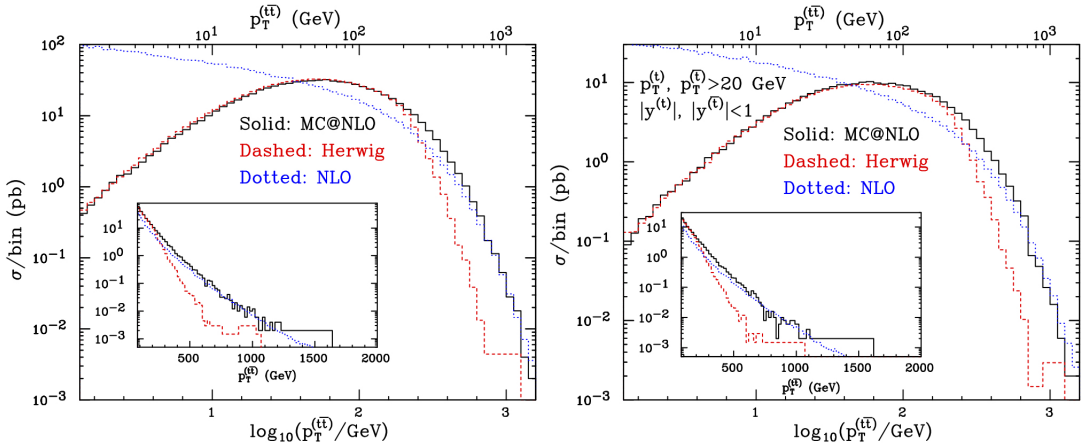


Fig. 58: The NLO+PS(MC@NLO) prediction for the transverse momentum of the top pair, $p_T^{(t\bar{t})}$, in top anti-top production at LHC14 with (right) and without (left) cuts on the top quarks, is compared with the PS(HERWIG) and the NLO fixed-order predictions. As expected, in the hight-p_T tail there is agreement between the NLO+PS and the NLO, while the PS prediction fails in this p_T region. On the other hand, at small p_T values the NLO+PS shows up the characteristic LL resummed result as given by the PS prediction.

cross section is much larger than the MC one, simply because hard emissions are correctly treated only in the former. For small transverse momentum of the top pair, the difference between the two histograms

shows the effect of all-order resummation, clearly, no meaningful comparison between NLO and data can be attempted in this region. In turn, we assist to a smooth interpolation performed by MC@NLO. For their precision, flexibility and availability the NLO+PS generators are the standard at the LHC.

A step further has been the combination of the two improvements mentioned above, having the merging and the matching at the same time. There have been several proposals on how to merge NLO computations with different jets in the final state. In general, one can think that there is always the need of a merging scale as the Q_1 we introduced discussing the CKKW algorithm. Nevertheless, for the class of processes with only one parton in the final state at the leading order (as H/W/Z+1jet) which, of course, presents divergences when this radiation becomes unresolved, it is possible to have finite results (even including the NLO corrections) just by multiplying the event weights by an extra Sudakov form factor that suppresses the divergences. It has been shown that such a Sudakov form factor can be chosen in such a way that the formal accuracy of the inclusive generation is again at Next-to-Leading Order [39]. This method is called MiNLO, that stands for Multi-scale improved NLO computation, and has the nice feature that it does not require a merging scale. As an example of application, we show the result for the associated production of an Higgs and a Vector boson at the LHC (VH with V=W or Z) [40]. In Fig.(59), the total cross section for proton proton collisions at 8 TeV is shown for both VH (in black) and VH plus one jet (VHJ) with MINLO (in red). When the usual scale variation bands are considered, we see that there is a good overlap of the two results.

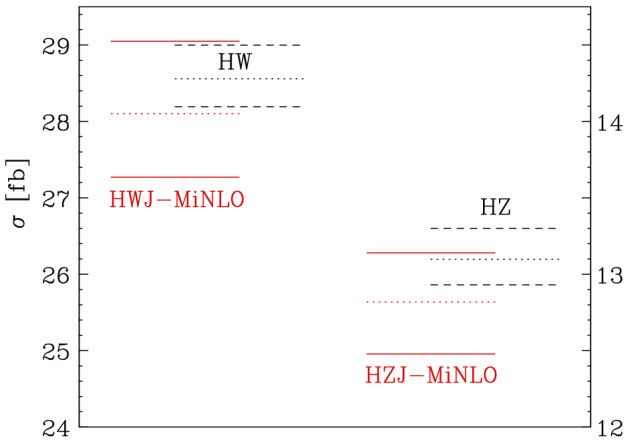


Fig. 59: Total cross section for associated Higgs production in proton-proton collision at LHC8. HV (V=W,Z) refers to the standard NLO computation, HVJ-MiNLO refers to the process with an extra jet treated with the MiNLO method.

In Fig.(60), we show the distribution of the transverse momentum of the WH system. For the high energy part of the distribution the two results obtained with WH and WHJ generators nicely overlap telling that the NLO correction provides only a small effect. In the low energy part of the distribution, instead, the two differ. The WHJ prediction, that contains the exact radiative correction for the first radiation, provides the most accurate result.

We conclude this section mentioning that reweighing a MiNLO NLO+PS generator for a process that has exactly one parton in the final state at the leading order (like H/W/Z+1jet, and in principle every process producing a colourless final state plus 1 jet) with a parton level NNLO computation for the fully exclusive production of such a colourless final state, one obtains an NNLO+PS matching! [41]. In Fig.(61), the Higgs boson transverse momentum distribution from gluon fusion at the LHC8 is shown.

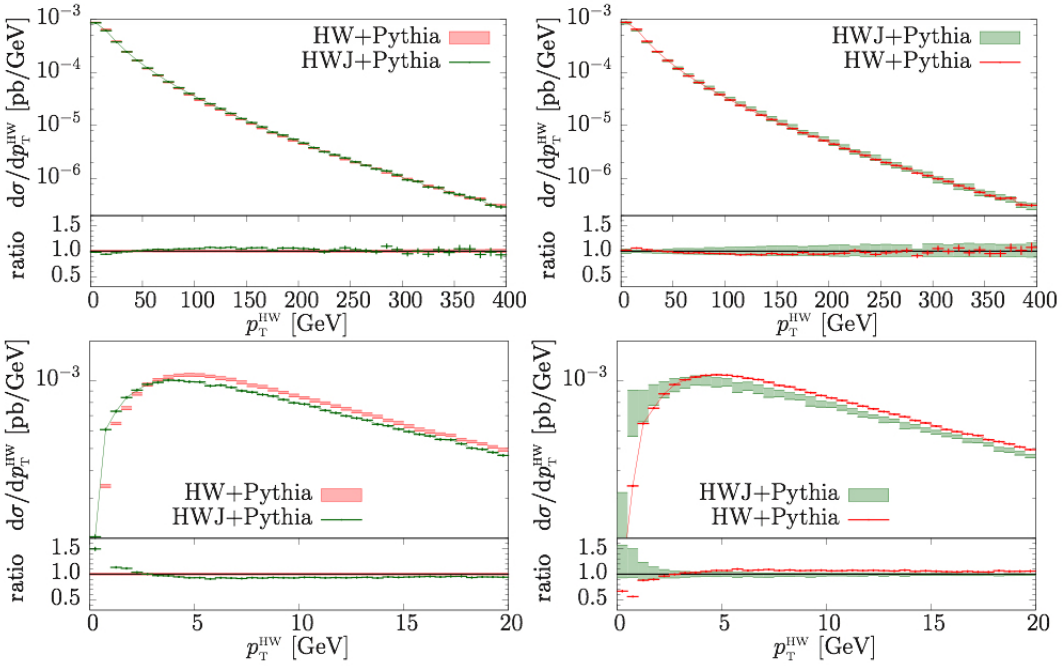


Fig. 60: Comparison between the HW and HWJ predictions for the p_T distribution of the HW system. In both case, generated events have been showered with PYTHIA. In the hight-p_T tail (upper panels), there is a good agreement while for small p_Ts, the HWJ-MiNLO predicts a slightly softer spectrum.

We remind that the fixed order result for this distribution (obtained with the program HNNLO [42] also

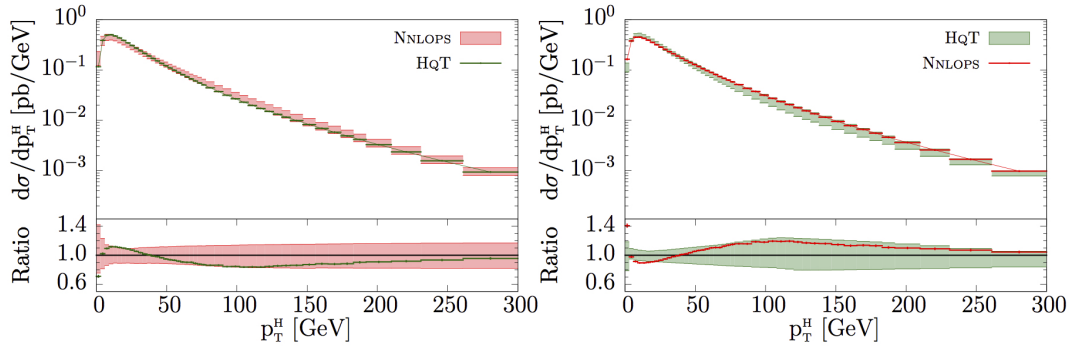


Fig. 61: Comparison between the NNLO+PS prediction and the resummed one for the p_T distribution of a Higgs boson produced in the gluon fusion channel at LHC8.

used to perform the reweight) is divergent in the low p_T region. The result obtained binning an event sample obtained with the NNLOPS generator is compared here to the result of the program HqT [43,44] that implements the resummation of the logarithmic enhancement discussed in a previous section. A reasonable agreement is found among the two when the scale variations are taken into account. Interestingly, the extension of the MiNLO and the NNLOPS techniques to the case with more partons in the final state is doable in principle, although it requires further non trivial analytic computations to build the appropriate Sudakov suppression factor, that are still missing.

Acknowledgements

I thank the organizers of the 2018 edition of the European Schools of High-Energy Physics for the excellent environment I've found at the school, all the discussion leaders for their great job in driving the discussion on the relevant points and all the students for their active and stimulating participation to the lectures. The discussion on several topics in these lectures have been strongly inspired by other lectures, in particular the ones by S. Catani held at the Galileo Galilei Institute in Italy[1] and by P. Nason at CERN[2], that I recommend.

References

[1] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev.*, D98(3):030001, 2018.

[2] Roel Aaij et al. Observation of the doubly charmed baryon Ξ_{cc}^{++}. *Phys. Rev. Lett.*, 119(11):112001, 2017.

[3] A. Esposito, A. L. Guerrieri, L. Maiani, F. Piccinini, A. Pilloni, A. D. Polosa, and V. Riquer. Observation of light nuclei at ALICE and the X(3872) conundrum. *Phys. Rev.*, D92(3):034028, 2015.

[4] R. Keith Ellis, W. James Stirling, and B. R. Webber. QCD and collider physics. *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.*, 8:1–435, 1996.

[5] Francisco J. Yndurain. *The Theory of Quark and Gluon Interactions*. Theoretical and Mathematical Physics. Springer, Berlin, Germany, 2006.

[6] P. A. Baikov, K. G. Chetyrkin, J. H. Kuhn, and J. Rittinger. Adler Function, Sum Rules and Crewther Relation of Order $O(\alpha_s^4)$: the Singlet Case. *Phys. Lett.*, B714:62–65, 2012.

[7] P. A. Baikov, K. G. Chetyrkin, and J. H. Kühn. Five-loop fermion anomalous dimension for a general gauge group from four-loop massless propagators. *JHEP*, 04:119, 2017.

[8] Thomas Luthe, Andreas Maier, Peter Marquard, and York Schröder. Five-loop quark mass and field anomalous dimensions for a general gauge group. *JHEP*, 01:081, 2017.

[9] F. Herzog, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt. The five-loop beta function of Yang-Mills theory with fermions. *JHEP*, 02:090, 2017.

[10] R. Keith Ellis, D. A. Ross, and A. E. Terrano. The Perturbative Calculation of Jet Structure in e+ e- Annihilation. *Nucl. Phys.*, B178:421–456, 1981.

[11] S. Catani and M. H. Seymour. A General algorithm for calculating jet cross-sections in NLO QCD. *Nucl. Phys.*, B485:291–419, 1997. [Erratum: Nucl. Phys.B510,503(1998)].

[12] S. Frixione, Z. Kunszt, and A. Signer. Three jet cross-sections to next-to-leading order. *Nucl. Phys.*, B467:399–442, 1996.

[13] W. Bartel et al. Experimental Studies on Multi-Jet Production in e+ e- Annihilation at PETRA Energies. *Z. Phys.*, C33:23, 1986. [,53(1986)].

[14] S. Catani, Yuri L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber. New clustering algorithm for multi - jet cross-sections in e+ e- annihilation. *Phys. Lett.*, B269:432–438, 1991.

[15] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys.*, B126:298–318, 1977.

[16] Richard D. Ball et al. Parton distributions with LHC data. *Nucl. Phys.*, B867:244–289, 2013.

[17] S. Alekhin, J. Blumlein, and S. Moch. Parton Distribution Functions and Benchmark Cross Sections at NNLO. *Phys. Rev.*, D86:054009, 2012.

[1] https : //www.youtube.com/playlist?list = PLDxsZU4NC6Z5ArDOepoRx4EfYmw6S_gne

[2] http : //cds.cern.ch/record/281780?

[18] P. Jimenez-Delgado and E. Reya. Dynamical NNLO parton distributions. *Phys. Rev.*, D79:074023, 2009.

[19] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J.*, C63:189–285, 2009.

[20] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne. Parton distributions in the LHC era: MMHT 2014 PDFs. *Eur. Phys. J.*, C75(5):204, 2015.

[21] Georges Aad et al. Measurement of inclusive two-particle angular correlations in pp collisions with the ATLAS detector at the LHC. *JHEP*, 05:157, 2012.

[22] Gavin P. Salam. Towards Jetography. *Eur. Phys. J.*, C67:637–686, 2010.

[23] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

[24] S. Catani, Yuri L. Dokshitzer, M. H. Seymour, and B. R. Webber. Longitudinally invariant K_t clustering algorithms for hadron hadron collisions. *Nucl. Phys.*, B406:187–224, 1993.

[25] Yuri L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber. Better jet clustering algorithms. *JHEP*, 08:001, 1997.

[26] Zoltan Nagy. Next-to-leading order calculation of three jet observables in hadron hadron collision. *Phys. Rev.*, D68:094002, 2003.

[27] Measurement of inclusive jet and dijet cross sections in proton-proton collision data at 7 TeV centre-of-mass energy using the ATLAS detector. 2011.

[28] James Currie, Aude Gehrmann-De Ridder, Thomas Gehrmann, E. W. N. Glover, Alexander Huss, and Joao Pires. Precise predictions for dijet production at the LHC. *Phys. Rev. Lett.*, 119(15):152001, 2017.

[29] Charalampos Anastasiou, Claude Duhr, Falko Dulat, Franz Herzog, and Bernhard Mistlberger. Higgs Boson Gluon-Fusion Production in QCD at Three Loops. *Phys. Rev. Lett.*, 114:212001, 2015.

[30] Bernhard Mistlberger. Higgs boson production at hadron colliders at N^3LO in QCD. *JHEP*, 05:028, 2018.

[31] Wojciech Bizoń, Xuan Chen, Aude Gehrmann-De Ridder, Thomas Gehrmann, Nigel Glover, Alexander Huss, Pier Francesco Monni, Emanuele Re, Luca Rottoli, and Paolo Torrielli. Fiducial distributions in Higgs and Drell-Yan production at N^3LL+NNLO. *JHEP*, 12:132, 2018.

[32] Torbjorn Sjostrand. Monte Carlo Generators. In *High-energy physics. Proceedings, European School, Aronsborg, Sweden, June 18-July 1, 2006*, pages 51–74, 2006.

[33] Stefan Höche. Introduction to parton-shower event generators. In *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014*, pages 235–295, 2015.

[34] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber. QCD matrix elements + parton showers. *JHEP*, 11:063, 2001.

[35] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, Roberto Pittau, and Antonio D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003.

[36] Stefano Frixione and Bryan R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 06:029, 2002.

[37] Paolo Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.

[38] Stefano Frixione, Paolo Nason, and Bryan R. Webber. Matching NLO QCD and parton showers in heavy flavor production. *JHEP*, 08:007, 2003.

[39] Keith Hamilton, Paolo Nason, Carlo Oleari, and Giulia Zanderighi. Merging H/W/Z + 0 and 1 jet

at NLO with no merging scale: a path to parton shower + NNLO matching. *JHEP*, 05:082, 2013.

[40] Gionata Luisoni, Paolo Nason, Carlo Oleari, and Francesco Tramontano. $HW^{\pm}$/HZ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO. *JHEP*, 10:083, 2013.

[41] Keith Hamilton, Paolo Nason, Emanuele Re, and Giulia Zanderighi. NNLOPS simulation of Higgs boson production. *JHEP*, 10:222, 2013.

[42] Stefano Catani and Massimiliano Grazzini. An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC. *Phys. Rev. Lett.*, 98:222002, 2007.

[43] Giuseppe Bozzi, Stefano Catani, Daniel de Florian, and Massimiliano Grazzini. Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC. *Nucl. Phys.*, B737:73–120, 2006.

[44] Daniel de Florian, Giancarlo Ferrera, Massimiliano Grazzini, and Damiano Tommasini. Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC. *JHEP*, 11:064, 2011.