QCD

Jan M. Pawlowski[1] and Tilman Plehn[1]

[1] *Institut für Theoretische Physik, Universität Heidelberg*

These lecture notes are the ambitious attempt to combine the two QCDs of the two lecturers into one coherent document. It includes many aspects of perturbative and non-perturbative QCD, including the basic theoretical concepts we need to compute observables at hadron colliders.

CONTENTS

I. Basics 3
A. Yang-Mills theory 3
B. QCD 6

II. Physics of Divergences 8
A. Ultraviolet divergences 8
B. Scaling logarithms 15
C. Infrared divergences 17
D. DGLAP equation 23
E. Collinear logarithms 29
F. Parton shower 35

III. Jets 42
A. Jet counting 42
B. Generating functional 44
C. Jets finders 51
D. Fat jets 53

I. BASICS

The theory of strong interactions, quantum chromodynamics (QCD), has been developed on the basis of scattering experiments that showed an internal $SU(3)$-symmetry and related charges much the same way quantum-electrodynamics (QED) shows the $U(1)$-symmetry related to the electric charge. The corresponding gauge theory, $SU(3)$ Yang-Mills theory, is non-Abelian and hence self-interacting, i.e. the (quantized) pure gauge theory is already non-trivial, in contrast to the $U(1)$-based QED.

A. Yang-Mills theory

We start by constructing the pure gauge part or Yang-Mills part of QCD as an $SU(3)$ gauge theory, fixing our conventions and repeating the main features known from the QFT II lecture. The weak $SU(2)$-theory turns out to have the same qualitative features as QCD (asymptotic freedom and confinement), but is technically simpler. On the other hand, the $SU(2)$ gauge bosons in the Standard Model are massive, leading to a major modification of this theory. Instead, we will assume massless gauge bosons throughout this lecture. As for QED, the classical action of QCD can be derived from the gauge-invariant (minimal) extension of the action of a free spin-one particle. The requirement of invariance of physics under local $SU(N_c)$ or color rotations with $\mathcal{U} \in SU(N_c)$, combined with a minimal coupling, leads us from partial to covariant derivatives,

$$\partial_\mu \to D_\mu(A) = \partial_\mu - i\, g\, A_\mu \,. \tag{I.1}$$

The gauge field A_μ in the adjoint representation is Lie-algebra–valued,

$$A_\mu = A_\mu^a\, t^a \,, \qquad \text{with} \qquad a = 1, ..., N_c^2 - 1 \,. \tag{I.2}$$

The matrices t^a are the generators of $SU(N_c)$. In physical QCD the gauge group has eight generators, $a = 1, ..., 8$, the Gell-Mann matrices. They are defined through

$$[t^a, t^b] = i\, f^{abc} t^c \,, \qquad\qquad \mathrm{tr_f}(t^a t^b) = \frac{1}{2}\delta^{ab} \,, \tag{I.3}$$

where the coefficients f^{abc} are the structure constants of the Lie algebra. and $\mathrm{tr_f}$ is the trace in the fundamental representation. The covariant derivative (I.1) does not carry any indices. In the adjoint representation it links to $SU(N_c)$ indices and reads

$$D_\mu^{ab}(A) = \partial_\mu \delta^{ab} - g\, f^{abc} A_\mu^c \,. \qquad \text{with} \qquad (t_{\rm ad}^c)^{ab} = -i\, f^{abc}. \tag{I.4}$$

The covariant derivative D_μ with its two color indices then has to transform as a tensor under gauge transformations,

$$D_\mu(A) \to D_\mu(A^{\mathcal{U}}) = \mathcal{U}\, D_\mu\, \mathcal{U}^\dagger \,, \qquad \text{with} \qquad \mathcal{U} = e^{i\omega} \in SU(N_c) \,, \tag{I.5}$$

where $\omega \in su(N_c)$ is the corresponding Lie algebra element. The covariance of D under gauge transformations in (I.5) implies

$$A_\mu \to A_\mu^{\mathcal{U}} = \frac{i}{g} U\, (D_\mu U^\dagger) = \mathcal{U}\, A_\mu\, \mathcal{U}^\dagger + \frac{i}{g} \mathcal{U}\, (\partial_\mu \mathcal{U}^\dagger) \,. \tag{I.6}$$

From the first term we confirm that in a non-Abelian gauge theory the gauge boson A_μ carries the corresponding color charge. There are various notations on the market leading to factors i and $-$ in the Lie algebra relations above. In the present lecture notes we have chosen hermitian generators which leads to the factor $+1/2$ for the trace in (I.3). It also entails real structure constants f^{abc} in the Lie-algebra in (I.3).

In analogy to QED the field strength tensor is defined through the commutator of covariant derivatives, it is the curvature tensor of the gauge theory. Based on the definitions in (I.1) and (I.3) we find

$$F_{\mu\nu} = \frac{i}{g}[D_\mu \,,\, D_\nu] = F_{\mu\nu}^a t^a \qquad \text{with} \qquad F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g\, f^{abc} A_\mu^b\, A_\nu^c. \tag{I.7}$$

Defined as in (I.7) the field strength $F_{\mu\nu}$ also transforms covariantly (as a tensor) under gauge transformations,

$$\begin{aligned} F_{\mu\nu}(A^{\mathcal{U}}) &= \frac{i}{g}\,[D_\mu(A^{\mathcal{U}}), D_\nu(A^{\mathcal{U}})] \\ &= \frac{i}{g}\,\mathcal{U}\,[D_\mu(A_\mu), D_\nu(A_\nu)]\,\mathcal{U}^\dagger = \mathcal{U}\,F_{\mu\nu}(A)\,\mathcal{U}^\dagger\,. \end{aligned} \tag{I.8}$$

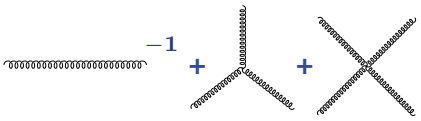
This allows us to define a gauge-invariant Yang-Mills (YM) action,

$$S_{\rm YM}[A] = \frac{1}{2}\int_x \mathrm{tr_f}\;(F_{\mu\nu}\,F_{\mu\nu}) = \frac{1}{4}\int_x F^a_{\mu\nu}\,F^a_{\mu\nu}\,, \tag{I.9}$$

with $\int_x = \int d^dx$. Its gauge invariance follows from (I.8),

$$S_{\rm YM}[A^{\mathcal{U}}] = \frac{1}{2}\int_x \mathrm{tr_f}\;\left(\mathcal{U}\,F_{\mu\nu}(A)F_{\mu\nu}(A)\,\mathcal{U}^\dagger\right) = S_{\rm YM}[A]\,, \tag{I.10}$$

where the last equality holds due to cyclicity of the trace in color space. Clearly, the action (I.9) with the field strength (I.7) is a self-interacting theory with coupling constant g. It has a quadratic kinetic term and three-gluon and four-gluon vertices. This is illustrated diagrammatically as

$S_{\rm YM}[A] \propto$ 

This allows us to read off the Feynman rules for the purely gluonic vertices. The full Feynman rules of QCD in the general covariant gauge are summarized in Fig. **??** in Appendix **??**. As in QED we can identify color-electric and color-magnetic fields as the components in the field strength tensor,

$$\begin{aligned} E^a_i &= F^a_{0i} \\ B^a_i &= \frac{1}{2}\epsilon_{ijk}F^a_{jk}. \end{aligned} \tag{I.11}$$

In contrast to QED these color-electric and magnetic fields are no observables, they change under gauge transformations. Only $\mathrm{tr}\,\vec{E}^2, \mathrm{tr}\,\vec{B}^2$ are observables.

A Yang-Mills theory can most easily be quantized through the path integral. Naively, the generating functional of pure YM-theory would read

$$Z[J] = \int dA\;\exp\left(-S_{\rm YM}[A] + \int_x J^a_\mu\,A^a_\mu\right). \tag{I.12}$$

The fundamental problem is that it contains redundant integrations due to gauge invariance of the action, see (I.10). These redundant integrations are usually removed by introducing a gauge fixing condition

$$\mathcal{F}[A_{\rm gf}] = 0 \tag{I.13}$$

Commonly used gauge fixings are

$$\begin{aligned} \partial_\mu A_\mu &= 0\,, \qquad \text{covariant or Lorenz gauge}\,, \\ \partial_i A_i &= 0\,, \qquad \text{Coulomb gauge}\,, \\ n_\mu A_\mu &= 0\,, \qquad \text{axial gauge}\,. \end{aligned} \tag{I.14}$$

The general covariant gauge has the technical advantage that it does not single out a space-time direction. This property reduces the possible tensor structure of correlation functions and hence simplifies computations. The Coulomb gauge and the axial gauge single out specific frames. At finite temperature (and density) this might be useful as the temperature singles out the thermal rest frame. In that case the Coulomb gauge and the temporal or Weyl gauge

($n_\mu = \delta_{\mu 0}$) are used often.

Gauge fields that are connected by gauge transformations are physically equivalent, i.e. their actions agree. They lie in so-called gauge orbits, $\{A^{\mathcal{U}}, \mathcal{U} \in SU(N)\}$, and fixing a gauge is equivalent to choosing a representative of such an orbit $A \to A_{\rm gf}$, up to potential (Gribov) copies. The occurrence of Gribov copies and how to handle them is discussed in Appendix **??**. To keep things simple we ignore them for the time being and continue with the construction of the QCD Lagrangian.

The path integral measure $\mathrm{d}A$ introduced in (I.12) can be split into an integration over physically inequivalent configurations $A_{\rm gf}$ and the gauge transformations $\mathcal{U}$,

$$\mathrm{d}A = J\,\mathrm{d}A_{\rm gf}\,\mathrm{d}\mathcal{U} \tag{I.15}$$

In (I.15) J denotes the Jacobian of the transformation $A \to (A_{\rm gf}, \mathcal{U})$, and we include $\mathrm{d}\mathcal{U}$ as the Haar measure of the gauge group, see e.g. [**?**]. The coordinate transformation (I.15) and the computation of the Jacobian J are done using the Faddeev-Popov quantization, [**?**]. To separate the integral (I.12) into the two parts shown in (I.15) we insert a very convoluted unity into the path integral,

$$1 = \int \mathrm{d}\mathcal{U}\,\delta\left[\mathcal{F}[A^{\mathcal{U}}]\right]\,\Delta_{\mathcal{F}}[A] = \Delta_{\mathcal{F}}[A]\int \mathrm{d}\mathcal{U}\,\delta\left[\mathcal{F}[A^{\mathcal{U}}]\right] \quad \Leftrightarrow \quad \Delta_{\mathcal{F}}[A] = \left(\int \mathrm{d}\mathcal{U}\,\delta\left[\mathcal{F}[A^{\mathcal{U}}]\right]\right)^{-1}, \tag{I.16}$$

where $\Delta_{\mathcal{F}}[A]$ is gauge-invariant due to the property $\mathrm{d}(\mathcal{U}\mathcal{V}) = \mathrm{d}\mathcal{U}$ of the Haar measure. For the path integral this gives us

$$\int \mathrm{d}A\; e^{-S_{\rm YM}[A]} = \int \mathrm{d}A\;\mathrm{d}\mathcal{U}\,\delta\left[\mathcal{F}[A^{\mathcal{U}}]\right]\,\Delta_{\mathcal{F}}[A]\; e^{-S_{\rm YM}[A]}\,. \tag{I.17}$$

Let us now consider a general observable $\mathcal{O}$, like *e.g.* $\mathrm{tr}F^2(x)\,\mathrm{tr}F^2(0)$. Observables are necessarily gauge invariant and local. The expectation value of $\mathcal{O}$ is defined as

$$\langle\mathcal{O}\rangle = \frac{\int \mathrm{d}A\,\mathcal{O}[A]\,e^{-S_{\rm YM}[A]}}{\int \mathrm{d}A\; e^{-S_{\rm YM}[A]}} = \frac{\int \mathrm{d}A\,\mathrm{d}\mathcal{U}\,\delta\left[\mathcal{F}[A^{\mathcal{U}}]\right]\,\Delta_{\mathcal{F}}[A]\,\mathcal{O}[A]\,e^{-S_{\rm YM}[A]}}{\int \mathrm{d}A\,\mathrm{d}\mathcal{U}\,\delta\left[\mathcal{F}[A^{\mathcal{U}}]\right]\,\Delta_{\mathcal{F}}[A]\,e^{-S_{\rm YM}[A]}}\,, \tag{I.18}$$

where we have simply inserted (I.16) into the path integral. In (I.18) all terms are gauge invariant except for the δ-function. Hence we can absorb the $\mathcal{U}$-dependence via $A \to A^{\mathcal{U}^\dagger}$. Then the (infinite) integral over the Haar measure decouples in numerator and denominator, and we arrive at

$$\langle\mathcal{O}\rangle = \frac{\int \mathrm{d}A\,\delta\left[\mathcal{F}[A]\right]\,\Delta_{\mathcal{F}}[A]\,\mathcal{O}[A]\,e^{-S_{\rm YM}[A_{\rm gf}]}}{\int \mathrm{d}A\,\delta\left[\mathcal{F}[A]\right]\,\Delta_{\mathcal{F}}[A]\,e^{-S_{\rm YM}[A_{\rm gf}]}}\,.$$

To compute the Jacobian $\Delta_{\mathcal{F}}[A]$ we apply a coordinate transformation to the δ-distribution

$$\delta[\mathcal{F}[A^{\mathcal{U}}]] = \frac{\delta[\omega-\omega_1]}{|\det\frac{\delta\mathcal{F}}{\delta\omega}|} \equiv \frac{\delta[\omega-\omega_1]}{|\det\mathcal{M}_{\mathcal{F}}[A]|} \quad \text{with} \quad \mathcal{U} = e^{i\omega}\,, \tag{I.19}$$

combined with a gauge fixing condition in the form (I.13)

$$\mathcal{F}[A_{\rm gf} = A^{\mathcal{U}(\omega_1)}] = 0\,. \tag{I.20}$$

Using the definition (I.16) this leads to

$$\Delta_{\mathcal{F}}[A] = |\det\mathcal{M}_{\mathcal{F}}[A_{\rm gf}]| \qquad \text{with} \qquad \mathcal{M}_{\mathcal{F}}[A] = \left.\frac{\delta\mathcal{F}}{\delta\omega}\right|_{\omega=0}[A]\,. \tag{I.21}$$

Here $A_{\rm gf}$ is the solution with the minimal distance to $A=0$. The inverse Jacobian $\det\mathcal{M}_{\mathcal{F}}$ of the ansatz (I.15) is called the Faddeev-Popov determinant. For its computation we consider an infinitesimal gauge transformation $\mathcal{U} = 1 + i\,g\,\omega$ where we have rescaled the transformation with the strong coupling g for convenience. Such a rescaling gives global factors of powers of $1/g$ that drop out in normalized expectation values. Then, the infinitesimal variation of the

covariant gauge $\partial_\mu A_\mu = 0$ follows as

$$\mathcal{F}[A^{\mathcal{U}}] = \partial_\mu A^{\mathcal{U}}_\mu = \partial_\mu A_\mu - \partial_\mu D_\mu \omega + O(\omega^2) \stackrel{!}{=} 0\,. \tag{I.22}$$

This gives us the Faddeev-Popov matrix

$$\mathcal{M}_{\mathcal{F}}[A] = -\frac{\delta \partial_\mu D_\mu \omega}{\delta \omega} = -\partial_\mu D_\mu \, \frac{\delta \omega}{\delta \omega} = -\partial_\mu D_\mu \mathbb{1}. \tag{I.23}$$

We assume that $-\partial^\mu D_\mu$ is a positive definite operator and we arrive at

$$\Delta_{\mathcal{F}}[A] = \det \mathcal{M}[A] = \det\left(-\partial_\mu D_\mu\right)\,. \tag{I.24}$$

A useful observation is that determinants can be represented by a Gaussian integral. In regular space such a Gaussian integral reads

$$\int_x e^{-\frac{1}{2}x^T M x} = \frac{(2\pi)^n}{\sqrt{\det M}}\,. \tag{I.25}$$

We want to use this relation to replace the Faddeev-Popov determinant (I.24) in the Lagrangian. It turns out that the usual form does not give a useful action or Lagrangian. However, we can instead use two anti-commuting Grassmann fields C and switch the sign in the exponent to

$$\det \mathcal{M}_{\mathcal{F}}[A] = \int \mathrm{d}c\,\mathrm{d}\bar{c}\, \exp\left\{\int d^dx\, d^dy\, \bar{c}^a(x) \mathcal{M}^{ab}_{\mathcal{F}}(x,y) c^b(y)\right\}\,. \tag{I.26}$$

Finally we slightly modify the gauge by introducing a Gaußian average over the gauges

$$\delta[\mathcal{F}[A^{\mathcal{U}}]] \to \int \mathrm{d}\mathcal{C}\, \delta[\mathcal{F}[A^{\mathcal{U}} - \mathcal{C}]] \exp\left\{-\frac{1}{2\xi}\int_x \mathcal{C}^a \mathcal{C}^a\right\}\,. \tag{I.27}$$

In summary, and restricting ourselves to the covariant gauge we then arrive at the generating functional for our Yang-Mills theory

$$Z[J_A, J_c, \bar{J}_c] = \int \mathrm{d}A\,\mathrm{d}c\,\mathrm{d}\bar{c}\, e^{-S_A[A,c,\bar{c}] + \int_x \left(J_A\cdot A + \bar{J}_c \cdot c - \bar{c}\cdot J_c\right)}\,. \tag{I.28}$$

The action including a general gauge fixing term and the Faddeev-Popov ghosts c^a is

$$S_A[A,c,\bar{c}] = \frac{1}{4}\int_x F^a_{\mu\nu} F^a_{\mu\nu} + \frac{1}{2\xi}\int_x \left(\partial_\mu A^a_\mu\right)^2 + \int_x \bar{c}^a \partial_\mu D^{ab}_\mu c^b\,, \tag{I.29}$$

where $\int_x = \int d^dx$ and the Landau gauge is achieved for $\xi = 0$. Note that the ghost action implies a negative dispersion for the ghost, related to the determinant of the positive operator $\mathcal{M}_{\mathcal{F}} = -\partial_\mu D_\mu$. However, this is a matter of convention, we might as well use a positive dispersion, the minus sign drops out for all correlation functions which do not involve ghosts, and only those are related to scattering amplitudes. The source term with all indices reads

$$\int_x \left(J_A \cdot A + \bar{J}_c \cdot c - \bar{c}\cdot J_c\right) \equiv \int_x \left(J^a_{A,\mu} A^a_\mu + \bar{J}^a_c c^a - \bar{c}^a J^a_c\right)\,. \tag{I.30}$$

The Feynman rules derived from (I.29) are summarized in Appendix **??**.

B. QCD

After briefly sketching the gauge part of QCD we now add fermionic matter fields. As before we start with the classical action, now given by the Dirac action of a quark doublet,

$$S_{\mathrm{Dirac}}[\psi, \bar{\psi}, A] = i\int_x \bar{\psi}\ \left(\not{D} + m_\psi + \mu\gamma_0\right)\ \psi\,, \tag{I.31}$$

where the Dirac matrices are defined through

$$\{\gamma_\mu\,,\,\gamma_\nu\} = 2\,\delta_{\mu\nu} \tag{I.32}$$

In (I.31), the fermions carry a Dirac index defining the 4-component spinor, gauge group indices in the fundamental representation of $SU(3)$, as well flavor indices. The latter we will ignore as long as we only talk about QCD and neglect the doublet nature of the matter fields in the Standard Model. The Dirac operator $\not{D}$ is diagonal in the flavor space as is the chemical potential term. The mass term depends on the current quark masses related to spontaneous symmetry breaking of the Higgs sector of the Standard Model. The up and down current quark masses are of the order $2-5$ MeV whereas the current quark mass of the strange quark is of the order 10^2 MeV. The other quark masses are of order $1-200$ GeV. In low energy QCD this has to be compared with the scale of strong chiral symmetry breaking $\Delta m \approx 300$ MeV. This mass scales are summarized in Table I.

Generation	first	second	third	Charge
Mass [MeV]	1.5-4	1150-1350	170×10^3	
Quark	u	c	t	$\frac{2}{3}$
Quark	d	s	b	$-\frac{1}{3}$
Mass [MeV]	4-8	80-130	$(4.1\text{-}4.4)\times10^3$	

TABLE I: Quark masses and charges. The scale of strong chiral symmetry breaking is $\Delta m \approx 300$ MeV as is $\Lambda_{\rm QCD}$. This entails that only $2+1$ flavours have to be considered for most applications to the phase diagram of QCD.

Evidently, for most applications of the QCD phase diagram we only have to consider the three lightest quark flavors, that is up, down and strange quark, to be dynamical. The current quark masses of up and down quarks are two order of magnitude smaller than all QCD infrared scales related to $\Lambda_{\rm QCD}$. Hence, the up and down quarks can be considered to be massless. This leads to the important observation that the physical masses of neutrons and protons — and hence the masses of the world around us — comes about from strong chiral symmetry breaking and has nothing to do with the Higgs sector.

In turn, the mass of the strange quark is of the order of $\Lambda_{\rm QCD}$ and has to be considered heavy for application in low energy QCD. The three heavier flavors, charm, bottom and top, are essentially static they do not contribute to the QCD dynamics relevant for its phase structure even though in particular the c-quark properties and bound states are much influenced by the infrared dynamics of QCD. In summary we will consider the $N_f = 2$ and $N_f = 2+1$ flavor cases for the phase structure of QCD, while for LHC physics all flavors are relevant.

Again in analogy to the Yang-Mills action we describe the quantized theory using its generating functional. The full generating functional of QCD is the straightforward extension of the Yang-Mills version in (I.28). The quark fields are Grassmann fields because of their fermionic nature and we are led to the generating functional

$$Z[J] = \int \mathrm{d}\Phi\, e^{-S_{\rm QCD}[\Phi]+\int_x\, J\cdot\phi}\,, \tag{I.33}$$

As a notation we have introduced super-fields and super-currents

$$\begin{aligned} \Phi &= (A, c, \bar c, \psi, \bar\psi) & J &= (J_A, J_c, \bar J_c, J_\psi, \bar J_\psi) \\ \mathrm{d}\Phi &= \int \mathrm{d}A\,\mathrm{d}c\,\mathrm{d}\bar c\,\mathrm{d}\psi\,\mathrm{d}\bar\psi & J\cdot\Phi &= J_A\cdot A + \bar J_c\cdot c - \bar c\cdot J_c + \bar J_\psi\cdot\psi - \bar\psi\cdot J_\psi\,. \end{aligned} \tag{I.34}$$

The gauge-fixed action $S_{\rm QCD}$ in (I.33) in the Landau gauge is given by

$$S_{\rm QCD}[\Phi] = \frac{1}{4}\int_x F^a_{\mu\nu}F^a_{\mu\nu} + \frac{1}{2\xi}\int_x \left(\partial_\mu A^a_\mu\right)^2 + \int_x \bar c^a \partial_\mu D^{ab}_\mu c^b + i\int_x \bar\psi\left(\not{D} + m_\psi + \mu\gamma_0\right)\psi\,. \tag{I.35}$$

The action in (I.35) is illustrated diagrammatically as
For physical observables the gauge dependence entering through the last two graphs in the first line, the ghost terms, is cancelled by the hidden gauge fixing dependence of the inverse gluon propagator. The Feynman rules are summarized in Appendix **??**.

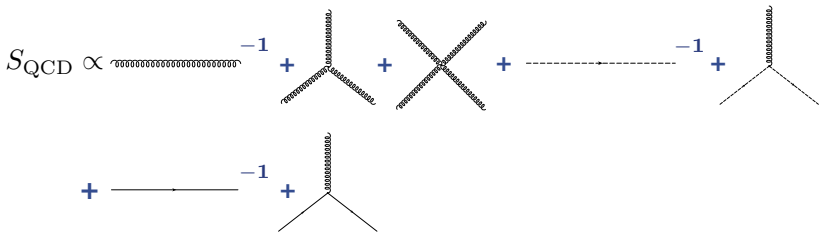


FIG. 1: Diagrammatical form of the QCD action.

II. PHYSICS OF DIVERGENCES

Now that we know the quantized action of QCD we can compute all kinds of processes to leading order in the strong couplings and beyond. From general field theory we know that when we are interested for example in cross section prediction with higher precision we need to compute further terms in its perturbative series in α_s. This computation will lead to ultraviolet divergences which can be absorbed into counter terms for any parameter in the Lagrangian. The crucial feature is that for a renormalizable theory like our Standard Model the number of counter terms is finite, which means once we know all parameters including their counter terms our theory becomes predictive.

We will see that in QCD processes we also encounter another kind of divergences. They arise from the infrared momentum regime. To understand their effects in LHC physics it is instructive to see what happens to the much better understood ultraviolet divergences. First, we will review how such ultraviolet divergences arise and how they are removed. Next, we will remind ourselves how running parameters appear in this procedure, *i.e.* how scale dependence is linked to the appearance of divergences. Finally, we need to interpret the use of running parameters physically and see that in perturbation theory they resum classes of logarithms to all orders in perturbation theory. For the infrared divergences we will follow exactly the same steps and develop some crucial features of hadron collider physics.

A. Ultraviolet divergences

Renormalization as the proper treatment of ultraviolet divergences is one of the most important things to understand about field theories; you can find more detailed discussions in any book on advanced field theory. The particular aspect of renormalization which will guide us through this section is the appearance of the renormalization scale.

In perturbation theory, scales automatically arise from the regularization of infrared or ultraviolet divergences. We can see this by writing down a simple scalar loop integral, with to two virtual scalar propagators with masses $m_{1,2}$ and an external momentum p flowing through a diagram

$$B(p^2; m_1, m_2) \equiv \int \frac{d^4q}{16\pi^2} \, \frac{1}{q^2 - m_1^2} \frac{1}{(q+p)^2 - m_2^2} \, . \tag{II.1}$$

Such two-point functions appear for example in the gluon self energy with virtual gluons, with massless ghost scalars, with a Dirac trace in the numerator for quarks, and with massive scalars for supersymmetric scalar quarks. In those cases the two masses are identical $m_1 = m_2$. The integration measure $1/(16\pi^2)$ is dictated by the Feynman rule for the integration over loop momenta. Counting powers of q in Eq.(II.1) we see that the integrand is not suppressed by powers of $1/q$ in the ultraviolet, so it is logarithmically divergent and we have to regularize it. Regularizing means expressing the divergence in a well–defined manner or scheme, allowing us to get rid of it by renormalization.

One regularization scheme is to introduce a cutoff into the momentum integral Λ, for example through the so-called Pauli—Villars regularization. Because the ultraviolet behavior of the integrand or integral cannot depend on any parameter living at a small energy scales, the parameterization of the ultraviolet divergence in Eq.(II.1) cannot involve the mass m or the external momentum p^2. The scalar two-point function has mass dimension zero, so its divergence has to be proportional to $\log(\Lambda/\mu_R)$ with a dimensionless prefactor and some scale μ_R^2 which is an artifact of the regularization of such a Feynman diagram.

A more elegant regularization scheme is dimensional regularization. It is designed not to break gauge invariance and naively seems to not introduce a mass scale μ_R. When we shift the momentum integration from 4 to $4 - 2\epsilon$ dimensions and use analytic continuation in the number of space–time dimensions to renormalize the theory, a

renormalization scale μ_R nevertheless appears once we ensure the two-point function and with it observables like cross sections keep their correct mass dimension

$$\int \frac{d^4q}{16\pi^2} \cdots \longrightarrow \mu_R^{2\epsilon} \int \frac{d^{4-2\epsilon}q}{16\pi^2} \cdots = \frac{i\mu_R^{2\epsilon}}{(4\pi)^2} \left[\frac{C_{-1}}{\epsilon} + C_0 + C_1\,\epsilon + \mathcal{O}(\epsilon^2) \right] . \tag{II.2}$$

At the end, the scale μ_R might become irrelevant and drop out after renormalization and analytic continuation, but to be on the safe side we keep it. The constants C_i in the series in $1/\epsilon$ depend on the loop integral we are considering. To regularize the ultraviolet divergence we assume $\epsilon > 0$ and find mathematically well defined poles $1/\epsilon$. Defining scalar integrals with the integration measure $1/(i\pi^2)$ will make for example C_{-1} come out as of the order $\mathcal{O}(1)$. This is the reason we usually find factors $1/(4\pi)^2 = \pi^2/(2\pi)^4$ in front of the loop integrals.

The poles in $1/\epsilon$ will cancel with the universal counter terms once we renormalize the theory. Counter terms we include by shifting parameters in the Lagrangian and the leading order matrix element. They cancel the poles in the combined leading order and virtual one-loop prediction

$$\begin{aligned} |\mathcal{M}_{\text{LO}}(g) + \mathcal{M}_{\text{virt}}|^2 &= |\mathcal{M}_{\text{LO}}(g)|^2 + 2\,\text{Re}\,\mathcal{M}_{\text{LO}}(g)\mathcal{M}_{\text{virt}} + \cdots \\ &\to |\mathcal{M}_{\text{LO}}(g+\delta g)|^2 + 2\,\text{Re}\,\mathcal{M}_{\text{LO}}(g)\mathcal{M}_{\text{virt}} + \cdots \\ \text{with} \qquad g \to g^{\text{bare}} &= g + \delta g \qquad \text{and} \quad \delta g \propto \alpha_s/\epsilon\,. \end{aligned} \tag{II.3}$$

The dots indicate higher orders in α_s, for example absorbing the δg corrections in the leading order and virtual interference. As we can see in Eq.(II.3) the counter terms do not come with a factor $\mu_R^{2\epsilon}$ in front. Therefore, while the poles $1/\epsilon$ cancel just fine, the scale factor $\mu_R^{2\epsilon}$ will not be matched between the actual ultraviolet divergence and the counter term.

We can keep track of the renormalization scale best by expanding the prefactor of the regularized but not yet renormalized integral in Eq.(II.2) in a Taylor series in ϵ, no question asked about convergence radii

$$\begin{aligned} \mu_R^{2\epsilon} \left[\frac{C_{-1}}{\epsilon} + C_0 + \mathcal{O}(\epsilon) \right] &= e^{2\epsilon \log \mu_R} \left[\frac{C_{-1}}{\epsilon} + C_0 + \mathcal{O}(\epsilon) \right] \\ &= \left[1 + 2\epsilon \log \mu_R + \mathcal{O}(\epsilon^2) \right] \left[\frac{C_{-1}}{\epsilon} + C_0 + \mathcal{O}(\epsilon) \right] \\ &= \frac{C_{-1}}{\epsilon} + C_0 + C_{-1}\,\log \mu_R^2 + \mathcal{O}(\epsilon) \\ &\to \frac{C_{-1}}{\epsilon} + C_0 + C_{-1}\,\log \frac{\mu_R^2}{M^2} + \mathcal{O}(\epsilon)\,. \end{aligned} \tag{II.4}$$

In the last step we correct by hand for the fact that $\log \mu_R^2$ with a mass dimension inside the logarithm cannot appear in our calculations. From somewhere else in our calculation the logarithm will be matched with a $\log M^2$ where M^2 is the typical mass or energy scale in our process. This little argument shows that also in dimensional regularization we introduce a mass scale μ_R which appears as $\log(\mu_R^2/M^2)$ in the renormalized expression for our observables. There is no way of removing ultraviolet divergences without introducing some kind of renormalization scale.

In Eq.(II.4) there appear two contributions to a given observable, the expected C_0 and the renormalization–induced C_{-1}. Because the factors C_{-1} are linked to the counter terms in the theory we can often guess them without actually computing the loop integral, which is very useful in cases where they numerically dominate.

Counter terms as they schematically appear in Eq.(II.3) are not uniquely defined. They need to include a given divergence to return finite observables, but we are free to add any finite contribution we want. This opens many ways to define a counter term for example based on physical processes where counter terms do not only cancel the pole but also finite contributions at a given order in perturbation theory. Needless to say, such schemes do not automatically work universally. An example for such a physical renormalization scheme is the on–shell scheme for masses, where we define a counter term such that external on–shell particles do not receive any corrections to their masses. For the top mass this means that we replace the leading order mass with the bare mass, for which we then insert the expression

in terms of the renormalized mass and the counter term

$$\begin{aligned} m_t^{\text{bare}} &= m_t + \delta m_t \\ &= m_t + m_t \frac{\alpha_s C_F}{4\pi} \left(3 \left(-\frac{1}{\epsilon} + \gamma_E - \log(4\pi) - \log \frac{\mu_R^2}{M^2} \right) - 4 + 3 \log \frac{m_t^2}{M^2} \right) \\ &\equiv m_t + m_t \frac{\alpha_s C_F}{4\pi} \left(-\frac{3}{\tilde{\epsilon}} - 4 + 3 \log \frac{m_t^2}{M^2} \right) \qquad \Leftrightarrow \qquad \frac{1}{\tilde{\epsilon}\left(\frac{\mu_R^2}{M^2}\right)} \equiv \frac{1}{\epsilon} - \gamma_E + \log \frac{4\pi\mu_R^2}{M^2} \,, \end{aligned} \tag{II.5}$$

with the color factor $C_F = (N^2-1)/(2N)$. The convenient scale dependent pole $1/\tilde{\epsilon}$ includes the universal additional terms like the Euler gamma function and the scaling logarithm. This logarithm is the big problem in this universality argument, since we need to introduce the at this stage arbitrary energy scale M to separate the universal logarithm of the renormalization scale and the parameter-dependent logarithm of the physical process.

A theoretical problem with this on–shell renormalization scheme is that it is not gauge invariant. On the other hand, it describes for example the kinematic features of top pair production at hadron colliders in a stable perturbation series. This means that once we define a more appropriate scheme for heavy particle masses in collider production mechanisms it better be numerically close to the pole mass. For the computation of total cross sections at hadron colliders or the production thresholds at e^+e^- colliders the pole mass is not well suited at all, but as we will see later this is not where we expect to measure particle masses at the LHC, so we should do fine with something very similar to the pole mass.

Another example for a process dependent renormalization scheme is the mixing of γ and Z propagators. There we choose the counter term of the weak mixing angle such that an on–shell Z boson cannot oscillate into a photon, and vice versa. We can generalize this scheme for mixing scalars as they for example appear in supersymmetry, but it is not gauge invariant with respect to the weak gauge symmetries of the Standard Model either. For QCD corrections, on the other hand, it is the most convenient scheme keeping all exchange symmetries of the two scalars.

To finalize this discussion of process dependent mass renormalization we quote the result for a scalar supersymmetric quark, a squark, where in the on–shell scheme we find

$$\begin{aligned} m_{\tilde{q}}^{\text{bare}} &= m_{\tilde{q}} + \delta m_{\tilde{q}} \\ &= m_{\tilde{q}} + m_{\tilde{q}} \frac{\alpha_s C_F}{4\pi} \left(-\frac{2r}{\tilde{\epsilon}} - 1 - 3r - (1-2r) \log r - (1-r)^2 \log \left| \frac{1}{r} - 1 \right| - 2r \log \frac{m_{\tilde{q}}^2}{M^2} \right) \,. \end{aligned} \tag{II.6}$$

with $r = m_{\tilde{g}}^2/m_{\tilde{q}}^2$. The interesting aspect of this squark mass counter term is that it also depends on the gluino mass, not just the squark mass itself. The reason why QCD counter terms tend to depend only on the renormalized quantity itself is that the gluon is massless. In the limit of vanishing gluino contribution the squark mass counter term is again only proportional to the squark mass itself

$$m_{\tilde{q}}^{\text{bare}} \Big|_{m_{\tilde{g}}=0} = m_{\tilde{q}} + \delta m_{\tilde{q}} = m_{\tilde{q}} + m_{\tilde{q}} \frac{\alpha_s C_F}{4\pi} \left(-\frac{1}{\tilde{\epsilon}} - 3 + \log \frac{m_{\tilde{q}}^2}{M^2} \right) \,. \tag{II.7}$$

Taking the limit of Eq.(II.6) to derive Eq.(II.7) is computationally not trivial, though.

One common feature of all mass counter terms listed above is $\delta m \propto m$, which means that we actually encounter a multiplicative renormalization

$$m^{\text{bare}} = Z_m \, m = (1 + \delta Z_m) \, m = \left(1 + \frac{\delta m}{m} \right) m = m + \delta m \,, \tag{II.8}$$

with $\delta Z_m = \delta m/m$ linking the two ways of writing the mass counter term. This form implies that particles with zero mass will not obtain a finite mass through renormalization. If we remember that chiral symmetry protects a Lagrangian from acquiring fermion masses this means that on–shell renormalization does not break this symmetry. A massless theory cannot become massive by mass renormalization. Regularization and renormalization schemes which do not break symmetries of the Lagrangian are ideal.

When we introduce counter terms in general field theory we usually choose a slightly more model independent

scheme — we define a renormalization point. This is the energy scale at which the counter terms cancels all higher order contributions, divergent as well as finite. The best known example is the electric charge which we renormalize in the Thomson limit of zero momentum transfer through the photon propagator

$$e \to e^{\text{bare}} = e + \delta e \; . \tag{II.9}$$

Looking back at δm_t as defined in Eq.(II.5) we also see a way to define a completely general counter term: if dimensional regularization, *i.e.* the introduction of $4 - 2\epsilon$ dimensions does not break any of the symmetries of our Lagrangian, like Lorentz symmetry or gauge symmetries, we can simply subtract the ultraviolet pole and nothing else. The only question is: do we subtract $1/\epsilon$ in the MS scheme or do we subtract $1/\tilde{\epsilon}$ in the $\overline{\text{MS}}$ scheme. In the $\overline{\text{MS}}$ scheme the counter term is then scale dependent.

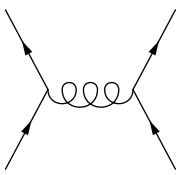
Carefully counting, there are three scales present in such a scheme. First, there is the physical scale in the process. In our case of a top self energy this is for example the top mass m_t appearing in the matrix element for the process $pp \to t\bar{t}$. Next, there is the renormalization scale μ_R, a reference scale which is part of the definition of any counter term. And last but not least, there is the scale M separating the counter term from the process dependent result, which we can choose however we want, but which as we will see implies a running of the counter term. The role of this scale M will become clear when we go through the example of the running strong coupling α_s. Of course, we would prefer to choose all three scales the same, but in a complex physical process this might not always be possible. For example, any massive $(2 \to 3)$ production process naturally involves several external physical scales.

Just a side remark for completeness: a one loop integral which has no intrinsic mass scale is the two-point function with zero mass in the loop and zero momentum flowing through the integral: $B(p^2 = 0; 0, 0)$. It appears for example in the self energy corrections of external quarks and gluons. Based on dimensional arguments this integral has to vanish altogether. On the other hand, we know that like any massive two-point function it has to be ultraviolet divergent $B \sim 1/\epsilon_{\text{UV}}$ because setting all internal and external mass scales to zero is nothing special from an ultraviolet point of view. This can only work if the scalar integral also has an infrared divergence appearing in dimensional regularization. We can then write the entire massless two-point function as

$$B(p^2 = 0; 0, 0) = \int \frac{d^4q}{16\pi^2} \, \frac{1}{q^2} \frac{1}{(q+p)^2} = \frac{i\pi^2}{16\pi^2} \left(\frac{1}{\epsilon_{\text{UV}}} - \frac{1}{\epsilon_{\text{IR}}} \right) \; , \tag{II.10}$$

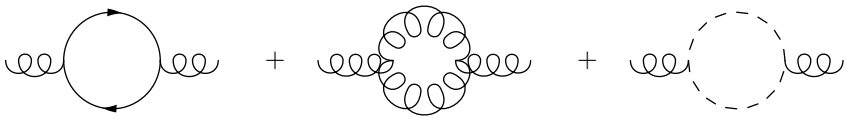
keeping track of the divergent contributions from the infrared and the ultraviolet regimes. For this particular integral they precisely cancel, so the result for $B(0; 0, 0)$ is zero, but setting it to zero too early will spoil any ultraviolet and infrared finiteness test. Treating the two divergences strictly separately and dealing with them one after the other also ensures that for ultraviolet divergences we can choose $\epsilon > 0$ while for infrared divergences we require $\epsilon < 0$.

To get an idea what these different scales which appear in the process of renormalization mean let us compute such a scale dependent parameter, namely the running strong coupling $\alpha_s(\mu_R^2)$. The Drell–Yan process is one of the very few relevant processes at hadron colliders where the strong coupling does not appear at tree level, so we cannot use it as our toy process this time. Another simple process where we can study this coupling is bottom pair production at the LHC, where at some energy range we will be dominated by valence quarks: $q\bar{q} \to b\bar{b}$. The only Feynman diagram is an s-channel off–shell gluon with a momentum flow $p^2 \equiv s$.



At next–to–leading order this gluon propagator will be corrected by self energy loops, where the gluon splits into two quarks or gluons and re-combines before it produces the two final–state bottoms. Let us for now assume that all quarks are massless. The Feynman diagrams for the gluon self energy include a quark look, a gluon loop, and the

ghost loop which removes the unphysical degrees of freedom of the gluon inside the loop.



The gluon self energy correction or vacuum polarization, as propagator corrections to gauge bosons are usually labelled, will be a scalar. This way, all fermion lines close in the Feynman diagram and the Dirac trace is computed inside the loop. In color space the self energy will (hopefully) be diagonal, just like the gluon propagator itself, so we can ignore the color indices for now. In unitary gauge the gluon propagator is proportional to the transverse tensor $T^{\mu\nu} = g^{\mu\nu} - p^\nu p^\mu/p^2$. As mentioned in the context of the effective gluon–Higgs coupling, the same should be true for the gluon self energy, which we therefore write as $\Pi^{\mu\nu} \equiv \Pi\, T^{\mu\nu}$. We find the simple relations

$$\begin{aligned} T^{\mu\nu} g^\rho_\nu &= \left(g^{\mu\nu} - \frac{p^\mu p^\nu}{p^2}\right) g^\rho_\nu = T^{\mu\rho} \\ T^{\mu\nu} T^\rho_\nu &= \left(g^{\mu\nu} - \frac{p^\mu p^\nu}{p^2}\right) \left(g^\rho_\nu - \frac{p_\nu p^\rho}{p^2}\right) = g^{\mu\rho} - 2\frac{p^\mu p^\rho}{p^2} + p^2 \frac{p^\mu p^\rho}{p^4} = T^{\mu\rho} \,. \end{aligned} \tag{II.11}$$

Including the gluon, quark, and ghost loops the regularized gluon self energy with a momentum flow p^2 through the propagator reads

$$\begin{aligned} -\frac{1}{p^2}\,\Pi\left(\frac{\mu_R^2}{p^2}\right) &= \frac{\alpha_s}{4\pi}\,\left(-\frac{1}{\tilde\epsilon(\mu_R^2/M^2)} + \log\frac{p^2}{M^2}\right)\,\left(\frac{13}{6}N_c - \frac{2}{3}n_f\right) + \mathcal{O}(\log m_t^2) \\ &\equiv \alpha_s\,\left(-\frac{1}{\tilde\epsilon(\mu_R^2/M^2)} + \log\frac{p^2}{M^2}\right)\,b_0 + \mathcal{O}(\log m_t^2) \\ &\text{with} \qquad \boxed{b_0 = \frac{1}{4\pi}\left(\frac{11}{3}N_c - \frac{2}{3}n_f\right) \overset{\text{SM}}{>} 0} \,. \end{aligned} \tag{II.12}$$

The minus sign arises from the factors i in the propagators. The number of fermions coupling to the gluons is n_f. From the comments on $B(p^2;0,0)$ we could guess that the loop integrals will only give a logarithm $\log p^2$ which is then matched by the logarithm $\log M^2$ implicitly included in the definition of $\tilde\epsilon$.

The factor b_0 arises from one-loop corrections, *i.e.* from diagrams which include one additional power of α_s. Strictly speaking, it gives the first term in a perturbative series in the strong coupling $\alpha_s = g_s^2/(4\pi)$. Later on, we will indicate where additional higher order corrections would enter.

In the second step of Eq.(II.12) we have sneaked in additional contributions to the renormalization of the strong coupling from the other one-loop diagrams in the process, replacing the factor 13/6 by a factor 11/3. This is related to the fact that there are actually three types of divergent virtual gluon diagrams in the physical process $q\bar q \to b\bar b$: the external quark self energies with renormalization factors $Z_f^{1/2}$, the internal gluon self energy Z_A, and the vertex corrections Z_{Aff}. The only physical parameters we can renormalize in this process are the strong coupling and, if finite, the bottom mass. Wave function renormalization constants are not physical, but vertex renormalization terms are. The entire divergence in our $q\bar q \to b\bar b$ process which needs to be absorbed in the strong coupling through Z_g is given by the combination

$$Z_{Aff} = Z_g Z_A^{1/2} Z_f \qquad\qquad \Leftrightarrow \qquad\qquad \frac{Z_{Aff}}{Z_A^{1/2} Z_f} \equiv Z_g \,. \tag{II.13}$$

We can check this definition of Z_g by comparing all vertices in which the strong coupling g_s appears, namely the gluon coupling to quarks, ghosts as well as the triple and quartic gluon vertex. All of them need to have the same divergence structure

$$\frac{Z_{Aff}}{Z_A^{1/2} Z_f} \overset{!}{=} \frac{Z_{A\eta\eta}}{Z_A^{1/2} Z_\eta} \overset{!}{=} \frac{Z_{3A}}{Z_A^{3/2}} \overset{!}{=} \sqrt{\frac{Z_{4A}}{Z_A^2}} \,. \tag{II.14}$$

If we had done the same calculation in QED and looked for a running electric charge, we would have found that the vacuum polarization diagrams for the photon do account for the entire counter term of the electric charge. The other two renormalization constants Z_{Aff} and Z_f cancel because of gauge invariance.

In contrast to QED, the strong coupling diverges in the Thomson limit because QCD is confined towards large distances and weakly coupled at small distances. Lacking a well enough motivated reference point we are lead to renormalize α_s in the $\overline{\text{MS}}$ scheme. From Eq.(II.12) we know that the ultraviolet pole which needs to be cancelled by the counter term is proportional to the function b_0

$$\begin{aligned}
& g_s^{\text{bare}} = Z_g g_s = (1+\delta Z_g)\, g_s = \left(1+\frac{\delta g_s}{g_s}\right) g_s \\
\Rightarrow \qquad & (g_s^2)^{\text{bare}} = (Z_g g_s)^2 = \left(1+\frac{\delta g_s}{g_s}\right)^2 g_s^2 = \left(1+2\frac{\delta g_s}{g_s}\right) g_s^2 = \left(1+\frac{\delta g_s^2}{g_s^2}\right) g_s^2 \\
\Rightarrow \qquad & \alpha_s^{\text{bare}} = \left(1+\frac{\delta\alpha_s}{\alpha_s}\right)\alpha_s \stackrel{!}{=} \left(1-\frac{\Pi}{p^2}\bigg|_{\text{pole}}\right)\alpha_s(M^2) \stackrel{\text{Eq.(II.12)}}{=} \left(1-\frac{\alpha_s}{\tilde{\epsilon}\left(\frac{\mu_R}{M}\right)}\, b_0\right)\alpha_s(M^2)\,. \qquad \text{(II.15)}
\end{aligned}$$

Only in the last step we have explicitly included the scale dependence of the counter term. Because the bare coupling does not depend on any scales, this means that α_s depends on the artifical external scale M. Similar to the top mass renormalization scheme we can switch to a more physical scheme for the strong coupling as well: we can absorb also the finite contributions of $\Pi(\mu_R^2/p^2)$ into the strong coupling by simply identifying $M^2 = p^2$. Based again on Eq.(II.12) this implies

$$\alpha_s^{\text{bare}} = \alpha_s(p^2)\left(1-\frac{\alpha_s(p^2)b_0}{\tilde{\epsilon}(\mu_R^2/M^2)} + \alpha_s(p^2)b_0\log\frac{p^2}{M^2}\right)\,. \qquad \text{(II.16)}$$

On the right hand side α_s is consistently evaluated as a function of the physical scale p^2. The lograrithm just shifts the argument of $\tilde{\epsilon}$ from M^2 to p^2. This formula defines a running coupling $\alpha_s(p^2)$, because the definition of the coupling now has to account for a possible shift between the original argument p^2 and the scale M^2 coming out of the $\overline{\text{MS}}$ scheme. Since according to Eqs.(II.15) and (II.16) the bare strong coupling can be expressed in terms of $\alpha_s(M^2)$ as well as in terms of $\alpha_s(p^2)$ we can link the two scales through

$$\begin{aligned}
& \alpha_s(M^2) = \alpha_s(p^2) + \alpha_s^2(p^2)b_0\log\frac{p^2}{M^2} = \alpha_s(p^2)\left(1+\alpha_s(p^2)b_0\log\frac{p^2}{M^2}\right) \\
\Leftrightarrow \qquad & \alpha_s(p^2) = \frac{\alpha_s(M^2)}{1+\alpha_s(p^2)b_0\log\frac{p^2}{M^2}} \\
\Leftrightarrow \qquad & \frac{d\alpha_s(p^2)}{d\log p^2} = -\alpha_s^2(p^2)b_0 + \mathcal{O}(\alpha_s^3)\,. \qquad \text{(II.17)}
\end{aligned}$$

To the given loop order the argument of the strong coupling squared in this formula can be neglected — its effect is of higher order. We nevertheless keep the argument as a higher order effect and remember the additional terms neglected above to later distinguish different approaches to the running coupling. From Eq.(II.12) we know that $b_0 > 0$, which means that towards larger scales the strong coupling has a negative slope. The ultraviolet limit of the strong coupling is zero. This makes QCD an asymptotically free theory. We can compute the function b_0 in general models by simply adding all contributions of strongly interacting particles in this loop

$$b_0 = -\frac{1}{12\pi}\sum_{\text{colored states}} D_j\, T_{R,j}\,, \qquad \text{(II.18)}$$

where we need to know some kind of counting factor D_j which is -11 for a vector boson (gluon), +4 for a Dirac fermion (quark), +2 for a Majorana fermion (gluino), +1 for a complex scalar (squark) and +1/2 for a real scalar. Note that this sign is not given by the fermionic or bosonic nature of the particle in the loop. The color charges are $T_R = 1/2$ for the fundamental representation of $SU(3)$ and $C_A = N_c$ for the adjoint representation. The masses of the loop particles are not relevant in this approximation because we are only interested in the ultraviolet regime of QCD where all particles can be regarded massless. This is a fundamental problem when we work with a running strong coupling constant: it is not an observable, which means that it does not have to ensure the decoupling of heavy states.

On the other hand, if we treat it like an observable we need to modify it by hand, so it does not ruin the automatic decoupling of heavy particles. When we really model the running of α_s we need to take into account threshold effects of heavy particles at their respective masses. When we really model the running of α_s we need to take into account threshold effects of heavy particles at their respective masses.

We can do even better than this fixed order in perturbation theory: while the correction to α_s in Eq.(II.16) is perturbatively suppressed by the usual factor $\alpha_s/(4\pi)$ it includes a logarithm of a ratio of scales which does not need to be small. Instead of simply including these gluon self energy corrections at a given order in perturbation theory we can instead include chains of one-loop diagrams with Π appearing many times in the off–shell gluon propagator. It means we replace the off–shell gluon propagator by

$$\begin{aligned}\frac{T^{\mu\nu}}{p^2} \to & \frac{T^{\mu\nu}}{p^2} + \left(\frac{T}{p^2}\cdot(-T\,\Pi)\cdot\frac{T}{p^2}\right)^{\mu\nu} \\ & + \left(\frac{T}{p^2}\cdot(-T\,\Pi)\cdot\frac{T}{p^2}\cdot(-T\,\Pi)\cdot\frac{T}{p^2}\right)^{\mu\nu} + \cdots \\ = & \frac{T^{\mu\nu}}{p^2}\sum_{j=0}^{\infty}\left(-\frac{\Pi}{p^2}\right)^j = \frac{T^{\mu\nu}}{p^2}\,\frac{1}{1+\Pi/p^2}\,, \end{aligned} \tag{II.19}$$

schematically written without the factors i. To avoid indices we abbreviate $T^{\mu\nu}T^{\rho}_{\nu} = T\cdot T$ which make sense because of $(T\cdot T\cdot T)^{\mu\nu} = T^{\mu\rho}T^{\sigma}_{\rho}T^{\nu}_{\sigma} = T^{\mu\nu}$. This resummation of the logarithm which appears in the next–to–leading order corrections to α_s moves the finite shift in α_s shown in Eqs.(II.12) and (II.16) into the denominator, while we assume that the pole will be properly taken care off in any of the schemes we discuss

$$\alpha_s^{\text{bare}} = \alpha_s(M^2) - \frac{\alpha_s^2 b_0}{\tilde{\epsilon}(\mu_R^2/M^2)} \equiv \frac{\alpha_s(p^2)}{1-\alpha_s(p^2)\,b_0\,\log\dfrac{p^2}{M^2}} - \frac{\alpha_s^2 b_0}{\tilde{\epsilon}(\mu_R^2/M^2)}\,. \tag{II.20}$$

Just as in the case without resummation, we can use this complete formula to relate the values of α_s at two reference points, *i.e.* we consider it a renormalization group equation (RGE) which evolves physical parameters from one scale to another in analogy to the fixed order version in Eq.(II.17)

$$\frac{1}{\alpha_s(M^2)} = \frac{1}{\alpha_s(p^2)}\left(1-\alpha_s(p^2)\,b_0\,\log\frac{p^2}{M^2}\right) = \frac{1}{\alpha_s(p^2)} - b_0\,\log\frac{p^2}{M^2}\,. \tag{II.21}$$

The factor α_s inside the parentheses we can again evaluate at either of the two scales, the difference is a higher order effect. If we keep it at p^2 we see that the expression in Eq.(II.21) is different from the un-resummed version in Eq.(II.16). If we ignore this higher order effect the two formulas become equivalent after switching p^2 and M^2. Resumming the vacuum expectation bubbles only differs from the un-resummed result once we include some next–to–leading order contribution. When we differentiate $\alpha_s(p^2)$ with respect to the momentum transfer p^2 we find, using the relation $d/dx(1/\alpha_s) = -1/\alpha_s^2\; d\alpha_s/dx$

$$\frac{1}{\alpha_s(p^2)}\,\frac{d\alpha_s(p^2)}{d\log p^2} = -\alpha_s(p^2)\,\frac{d}{d\log p^2}\,\frac{1}{\alpha_s(p^2)} = -\alpha_s(p^2)\,b_0 \qquad \text{or} \qquad \boxed{p^2\frac{d\alpha_s}{dp^2} \equiv \frac{d\alpha_s}{d\log p^2} = \beta = -\alpha_s^2\sum_{n=0} b_n\alpha_s^n}\,. \tag{II.22}$$

In the second form we replace the one-loop running b_0 by its full perturbative series. This is the famous running of the strong coupling constant including all higher order terms b_n.

In the running of the strong coupling constant we relate the different values of α_s through multiplicative factors of the kind

$$\left(1 \pm \alpha_s(p^2) b_0 \log\frac{p^2}{M^2}\right)\,. \tag{II.23}$$

Such factors appear in the un-resummed computation of Eq.(II.17) as well as in Eq.(II.20) after resummation. Because they are multiplicative, these factors can move into the denominator, where we need to ensure that they do not vanish. Dependent on the sign of b_0 this becomes a problem for large scale ratios $|\alpha_s \log p^2/M^2| > 1$, where it leads to the

Landau pole. For the strong coupling with $b_0 > 0$ and large coupling values at small scales $p^2 \ll M^2$ the combination $(1 + \alpha_s b_0 \log p^2/M^2)$ can indeed vanish and become a problem.

It is customary to replace the renormalization point of α_s in Eq.(II.20) with a reference scale defined by the Landau pole. At one loop order this reads

$$\begin{aligned} 1 + \alpha_s\, b_0\, \log \frac{\Lambda^2_{\rm QCD}}{M^2} \stackrel{!}{=} 0 \qquad &\Leftrightarrow \qquad \log \frac{\Lambda^2_{\rm QCD}}{M^2} = -\frac{1}{\alpha_s(M^2) b_0} \qquad \Leftrightarrow \qquad \log \frac{p^2}{M^2} = \log \frac{p^2}{\Lambda^2_{\rm QCD}} - \frac{1}{\alpha_s(M^2) b_0} \\ \frac{1}{\alpha_s(p^2)} &\stackrel{\text{Eq.(II.21)}}{=} \frac{1}{\alpha_s(M^2)} + b_0\, \log \frac{p^2}{M^2} \\ &= \frac{1}{\alpha_s(M^2)} + b_0 \log \frac{p^2}{\Lambda^2_{\rm QCD}} - \frac{1}{\alpha_s(M^2)} = b_0 \log \frac{p^2}{\Lambda^2_{\rm QCD}} \qquad \Leftrightarrow \qquad \boxed{\alpha_s(p^2) = \frac{1}{b_0 \log \dfrac{p^2}{\Lambda^2_{\rm QCD}}}} \,. \end{aligned} \tag{II.24}$$

This scheme can be generalized to any order in perturbative QCD and is not that different from the Thomson limit renormalization scheme of QED, except that with the introduction of $\Lambda_{\rm QCD}$ we are choosing a reference point which is particularly hard to compute perturbatively. One thing that is interesting in the way we introduce $\Lambda_{\rm QCD}$ is the fact that we introduce a scale into our theory without ever setting it. All we did was renormalize a coupling which becomes strong at large energies and search for the mass scale of this strong interaction. This trick is called dimensional transmutation.

In terms of language, there is a little bit of confusion between field theorists and phenomenologists: up to now we have introduced the renormalization scale μ_R as the renormalization point, for example of the strong coupling constant. In the $\overline{\rm MS}$ scheme, the subtraction of $1/\tilde{\epsilon}$ shifts the scale dependence of the strong coupling to M^2 and moves the logarithm $\log M^2/\Lambda^2_{\rm QCD}$ into the definition of the renormalized parameter. This is what we will from now on call the renormalization scale in the phenomenological sense, *i.e.* the argument we evaluate α_s at. Throughout this section we will keep the symbol M for this renormalization scale in the $\overline{\rm MS}$ scheme, but from Section II C on we will shift back to μ_R instead of M as the argument of the running coupling, to be consistent with the literature.

B. Scaling logarithms

In the last section we have introduced the running strong coupling in a fairly abstract manner. For example, we did not link the resummation of diagrams and the running of α_s in Eqs.(II.17) and (II.22) to physics. In what way does the resummation of the one-loop diagrams for the s-channel gluon improve our prediction of an observable? To illustrate this we best look at a simple observable which depends on just one energy scale p^2. The first observable coming to mind is again the Drell–Yan cross section $\sigma(q\bar{q} \to \mu^+\mu^-)$, but since we are not really sure what to do with the parton densities which are included in the actual hadronic observable, we better use an observable at an e^+e^- collider. Something that will work and includes α_s at least in the one-loop corrections is the R parameter

$$R = \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} = N_c \sum_{\text{quarks}} Q_q^2 = \frac{11 N_c}{9} \,. \tag{II.25}$$

The numerical value at leading order assumes five quarks. Including higher order corrections we can express the result in a power series in the renormalized strong coupling α_s. In the $\overline{\rm MS}$ scheme we subtract $1/\tilde{\epsilon}(\mu_R^2/M^2)$ and in general include a scale dependence on M in the individual prefactors r_n

$$R\left(\frac{p^2}{M^2}, \alpha_s\right) = \sum_{n=0} r_n\left(\frac{p^2}{M^2}\right)\, \alpha_s^n(M^2) \qquad\qquad r_0 = \frac{11 N_c}{9} \,. \tag{II.26}$$

The r_n we can assume to be dimensionless — if they are not, we can scale R appropriately using p^2. This implies that the r_n only depend on ratios of two scales, the externally fixed p^2 on the one hand and the artificial M^2 on the other.

At the same time we know that R is an observable, which means that including all orders in perturbation theory it cannot depend on any artificial scale choice M. Writing this dependence as a total derivative and setting it to zero

we find an equation which would be called a Callan–Symanzik equation if instead of the running coupling we had included a running mass

$$\begin{aligned}
0 \stackrel{!}{=} M^2 \frac{d}{dM^2} R\left(\frac{p^2}{M^2}, \alpha_s(M^2)\right) &= M^2 \left[\frac{\partial}{\partial M^2} + \frac{\partial \alpha_s}{\partial M^2}\frac{\partial}{\partial \alpha_s}\right] R\left(\frac{p^2}{M^2}, \alpha_s\right) \\
&= \left[M^2 \frac{\partial}{\partial M^2} + \beta \frac{\partial}{\partial \alpha_s}\right] \sum_{n=0} r_n \left(\frac{p^2}{M^2}\right) \alpha_s^n \\
&= \sum_{n=1} M^2 \frac{\partial r_n}{\partial M^2} \alpha_s^n + \sum_{n=1} \beta\, r_n\, n \alpha_s^{n-1} \qquad \text{with} \quad r_0 = \frac{11 N_c}{9} = \text{const} \\
&= M^2 \sum_{n=1} \frac{\partial r_n}{\partial M^2} \alpha_s^n - \sum_{n=1} \sum_{m=0} n r_n\, \alpha_s^{n+m+1}\, b_m \qquad \text{with} \quad \beta = -\alpha_s^2 \sum_{m=0} b_m \alpha_s^m \\
&= M^2 \frac{\partial r_1}{\partial M^2} \alpha_s + \left(M^2 \frac{\partial r_2}{\partial M^2} - r_1 b_0\right) \alpha_s^2 + \left(M^2 \frac{\partial r_3}{\partial M^2} - 2 r_2 b_0 - r_1 b_1\right) \alpha_s^3 + \mathcal{O}(\alpha_s^4) \,. \qquad \text{(II.27)}
\end{aligned}$$

In the second line we have to remember that the M dependence of α_s is already included in the appearance of β, so α_s should be considered a variable by itself. This perturbative series in α_s has to vanish in each order of perturbation theory. The non–trivial structure, namely the mix of r_n derivatives and the perturbative terms in the β function we can read off the α_s^3 term in Eq.(II.27): first, we have the appropriate NNNLO corrections r_3. Next, we have one loop in the gluon propagator b_0 and two loops for example in the vertex r_2. And finally, we need the two-loop diagram for the gluon propagator b_1 and a one-loop vertex correction r_1. The kind–of–Callan–Symanzik equation Eq.(II.27) requires

$$\begin{aligned}
\frac{\partial r_1}{\partial \log M^2/p^2} &= 0 \\
\frac{\partial r_2}{\partial \log M^2/p^2} &= r_1 b_0 \\
\frac{\partial r_3}{\partial \log M^2/p^2} &= r_1 b_1 + 2 r_2(M^2) b_0 \\
&\dots \qquad \text{(II.28)}
\end{aligned}$$

The dependence on the argument M^2 vanishes for r_0 and r_1. Keeping in mind that there will be integration constants c_n and that another, in our case, unique momentum scale p^2 has to cancel the mass units inside $\log M^2$ we find

$$\begin{aligned}
r_0 &= c_0 = \frac{11 N_c}{9} \\
r_1 &= c_1 \\
r_2 &= c_2 + r_1 b_0 \log \frac{M^2}{p^2} = c_2 + c_1 b_0 \log \frac{M^2}{p^2} \\
r_3 &= \int d \log \frac{M'^2}{p^2} \left(c_1 b_1 + 2\left(c_2 + c_1 b_0 \log \frac{M'^2}{p^2}\right) b_0\right) \\
&= c_3 + (c_1 b_1 + 2 c_2 b_0) \log \frac{M^2}{p^2} + c_1 b_0^2 \log^2 \frac{M^2}{p^2} \\
&\dots \qquad \text{(II.29)}
\end{aligned}$$

This chain of r_n values looks like we should interpret the apparent fixed-order perturbative series for R in Eq.(II.26) as a series which implicitly includes terms of the order $\log^{n-1} M^2/p^2$ in each r_n. They can become problematic if this logarithm becomes large enough to spoil the fast convergence in terms of $\alpha_s \sim 0.1$, evaluating the observable R at scales far away from the scale choice for the strong coupling constant M.

Instead of the series in r_n we can use the conditions in Eq.(II.29) to express R in terms of the c_n and collect the

logarithms appearing with each c_n. The geometric series we then resum to

$$\begin{aligned} R = \sum_n r_n \left(\frac{p^2}{M^2}\right) \alpha_s^n(M^2) &= c_0 + c_1 \left(1 + \alpha_s(M^2) b_0 \log \frac{M^2}{p^2} + \alpha_s^2(M^2) b_0^2 \log^2 \frac{M^2}{p^2} + \dots\right) \alpha_s(M^2) \\ &\quad + c_2 \left(1 + 2\alpha_s(M^2) b_0 \log \frac{M^2}{p^2} + \dots\right) \alpha_s^2(M^2) + \dots \\ &= c_0 + c_1 \frac{\alpha_s(M^2)}{1 - \alpha_s(M^2) b_0 \log \frac{M^2}{p^2}} + c_2 \left(\frac{\alpha_s(M^2)}{1 - \alpha_s(M^2) b_0 \log \frac{M^2}{p^2}}\right)^2 + \dots \\ &\equiv \sum c_n \, \alpha_s^n(p^2) \, . \end{aligned} \tag{II.30}$$

In the original ansatz α_s is always evaluated at the scale M^2. In the last step we use Eq.(II.21) with flipped arguments p^2 and M^2, derived from the resummation of the vacuum polarization bubbles. In contrast to the r_n integration constants the c_n are by definition independent of p^2/M^2 and therefore more suitable as a perturbative series in the presence of potentially large logarithms. Note that the un-resummed version of the running coupling in Eq.(II.16) would not give the correct result, so Eq.(II.30) only holds for resummed vacuum polarization bubbles.

This re-organization of the perturbation series for R can be interpreted as resumming all logarithms of the kind $\log M^2/p^2$ in the new organization of the perturbative series and absorbing them into the running strong coupling evaluated at the scale p^2. All scale dependence in the perturbative series for the dimensionless observable R is moved into α_s, so possibly large logarithms $\log M^2/p^2$ have disappeared. In Eq.(II.30) we also see that this series in c_n will never lead to a scale-invariant result when we include a finite order in perturbation theory. Some higher–order factors c_n are known, for example inserting $N_c = 3$ and five quark flavors just as we assume in Eq.(II.25)

$$R = \frac{11}{3} \left(1 + \frac{\alpha_s(p^2)}{\pi} + 1.4 \left(\frac{\alpha_s(p^2)}{\pi}\right)^2 - 12 \left(\frac{\alpha_s(p^2)}{\pi}\right)^3 + \mathcal{O}\left(\frac{\alpha_s(p^2)}{\pi}\right)^4\right) . \tag{II.31}$$

This alternating series with increasing perturbative prefactors seems to indicate the asymptotic instead of convergent behavior of perturbative QCD. At the bottom mass scale the relevant coupling factor is only $\alpha_s(m_b^2)/\pi \sim 1/14$, so a further increase of the c_n would become dangerous. However, a detailed look into the calculation shows that the dominant contributions to c_n arise from the analytic continuation of logarithms, which are large finite terms for example from $\text{Re}(\log^2(-E^2)) = \log^2 E^2 + \pi^2$. In the literature such π^2 terms arising from the analytic continuation of loop integrals are often phrased in terms of $\zeta_2 = \pi^2/6$.

Before moving on we collect the logic of the argument given in this section: when we regularize an ultraviolet divergence we automatically introduce a reference scale μ_R. Naively, this could be an ultraviolet cutoff scale, but even the seemingly scale invariant dimensional regularization in the conformal limit of our field theory cannot avoid the introduction of a scale. There are several ways of dealing with such a scale: first, we can renormalize our parameter at a reference point. Secondly, we can define a running parameter and this way absorb the scale logarithm into the $\overline{\text{MS}}$ counter term. In that case introducing Λ_{QCD} leaves us with a compact form of the running coupling $\alpha_s(M^2; \Lambda_{\text{QCD}})$.

Strictly speaking, at each order in perturbation theory the scale dependence should vanish together with the ultraviolet poles, as long as there is only one scale affecting a given observable. However, defining the running strong coupling we sum one-loop vacuum polarization graphs. Even when we compute an observable at a given loop order, we implicitly include higher order contributions. They lead to a dependence of our perturbative result on the artificial scale M^2, which phenomenologists refer to as renormalization scale dependence.

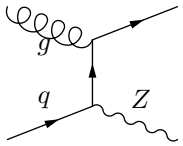
Using the R ratio we see what our definition of the running coupling means in terms of resumming logarithms: reorganizing our perturbative series to get rid of the ultraviolet divergence $\alpha_s(p^2)$ resums the scale logarithms $\log p^2/M^2$ to all orders in perturbation theory. We will need this picture once we introduce infrared divergences in the following section.

C. Infrared divergences

After our brief discussion into ultraviolet divergences and renormalization we move to LHC processes, like the Drell–Yan process $q\bar{q} \to \mu^+\mu^-$, in perturbative QCD. We know that the partons inside the proton are described by parton

distributions (pdfs), which at this stage are only probability functions in terms of the collinear momentum fraction of the partons inside the proton. The question we need to ask for a quantum-level description of $\mu^+\mu^-$ production at the LHC is: what happens if together with the two leptons we produce additional jets which for one reason or another we do not observe in the detector. Such jets could for example come from the radiation of a gluon from the initial–state quarks. We will first study the kinematics of radiating such jets and specify the infrared divergences this leads to. Next we will show that these divergences have a generic structure and can be absorbed into a re-definition of the parton densities, similar to an ultraviolet renormalization of a Lagrangian parameter. Finally, we will follow the example of the ultraviolet divergences and see what absorbing these divergences means in terms logarithms appearing in QCD calculations.

Let us get back to the radiation of additional partons in the Drell–Yan process. We can start for example by computing the cross section for the partonic process $q\bar{q} \to Zg$. However, this partonic process involves renormalization of ultraviolet divergences as well as loop diagrams which we have to include before we can say anything reasonable. Instead, we look at the crossed process



It should behave similar to any other $(2 \to 2)$ jet radiation, except that it has a different incoming state than the leading order Drell–Yan process and hence does not involve virtual corrections. This means we do not have to deal with ultraviolet divergences and renormalization, and can concentrate on parton or jet radiation from the initial state. Moreover, let us go back to Z production instead of a photon, to avoid confusion with additional massless particles in the final state.

The amplitude for this $(2 \to 2)$ process is — modulo charges and averaging factors, but including all Mandelstam variables

$$|\mathcal{M}|^2 \sim -\frac{t}{s} - \frac{s^2 - 2m_Z^2(s+t-m_Z^2)}{st} \; . \tag{II.32}$$

The Mandelstam variable t for one massless final–state particle can be expressed as $t = -s(1-\tau)y$ in terms of the rescaled gluon emission angle $y = (1-\cos\theta)/2$ and $\tau = m_Z^2/s$. Similarly, we obtain $u = -s(1-\tau)(1-y)$, so as a first check we can confirm that $t + u = -s(1-\tau) = -s + m_Z^2$. The collinear limit when the gluon is radiated in the beam direction is given by $y \to 0$, corresponding to negative $t \to 0$ with finite $u = -s + m_Z^2$. In this limit the matrix element can also be written as

$$|\mathcal{M}|^2 \sim \frac{s^2 - 2sm_Z^2 + 2m_Z^4}{s(s-m_Z^2)} \; \frac{1}{y} + \mathcal{O}(y^0) \; . \tag{II.33}$$

This expression is divergent for collinear gluon radiation or gluon splitting, *i.e.* for small angles y. We can translate this $1/y$ divergence for example into the transverse momentum of the gluon or Z

$$sp_T^2 = tu = s^2(1-\tau)^2 \; y(1-y) = (s-m_Z^2)^2 y + \mathcal{O}(y^2) \tag{II.34}$$

In the collinear limit our matrix element squared in Eq.(II.33) becomes

$$\boxed{|\mathcal{M}|^2 \sim \frac{s^2 - 2sm_Z^2 + 2m_Z^4}{s^2} \; \frac{s-m_Z^2}{p_T^2} + \mathcal{O}(p_T^0)} \; . \tag{II.35}$$

The matrix element for the tree level process $qg \to Zq$ has a leading divergence proportional to $1/p_T^2$. To compute the total cross section for this process we need to integrate the matrix element over the entire two-particle phase space. Using the appropriate Jacobian this integration can be written in terms of the reduced angle y. Approximating the matrix element as C'/y or C/p_T^2, we then integrate

$$\int_{y^{\min}}^{y^{\max}} dy \frac{C'}{y} = \int_{p_T^{\min}}^{p_T^{\max}} dp_T^2 \frac{C}{p_T^2} = \; 2 \int_{p_T^{\min}}^{p_T^{\max}} dp_T \; p_T \; \frac{C}{p_T^2} \simeq 2C \int_{p_T^{\min}}^{p_T^{\max}} dp_T \frac{1}{p_T} = 2C \; \log \frac{p_T^{\max}}{p_T^{\min}} \tag{II.36}$$

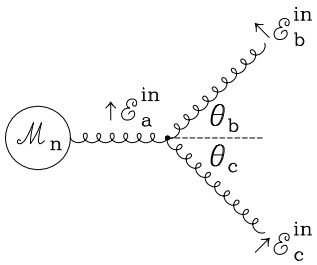


FIG. 2: Splitting of one gluon into two gluons. Figure from Ref. [**?**].

The form C/p_T^2 for the matrix element is of course only valid in the collinear limit; in the non–collinear phase space C is not a constant. However, Eq.(II.36) describes well the collinear divergence arising from quark radiation at the LHC.

We now follow the same strategy as for the ultraviolet divergence. First, we regularize the divergence for example using dimensional regularization. Then, we find a well–defined way to get rid of it. Dimensional regularization means writing the two-particle phase space in $n = 4 - 2\epsilon$ dimensions. Just for reference, the complete formula in terms of the angular variable y then reads

$$s\,\frac{d\sigma}{dy} = \frac{\pi(4\pi)^{-2+\epsilon}}{\Gamma(1-\epsilon)}\,\left(\frac{\mu_F^2}{m_Z^2}\right)^\epsilon\,\frac{\tau^\epsilon(1-\tau)^{1-2\epsilon}}{y^\epsilon(1-y)^\epsilon}\,|\mathcal{M}|^2 \sim \left(\frac{\mu_F^2}{m_Z^2}\right)^\epsilon\,\frac{|\mathcal{M}|^2}{y^\epsilon(1-y)^\epsilon}\,. \tag{II.37}$$

In the second step we only keep the factors we are interested in. The additional factor $1/y^\epsilon$ regularizes the integral at $y \to 0$, as long as $\epsilon < 0$ by slightly increasing the suppression of the integrand in the infrared regime. This means that for infrared divergences we can as well choose $n = 4 + 2\epsilon$ space–time dimensions with $\epsilon > 0$. After integrating the leading collinear divergence $1/y^{1+\epsilon}$ we are left with a pole $1/(-\epsilon)$. This regularization procedure is symmetric in $y \leftrightarrow (1-y)$. What is important to notice is again the appearance of a scale $\mu_F^{2\epsilon}$ with the n-dimensional integral. This scale arises from the infrared regularization of the phase space integral and is referred to as factorization scale. The actual removal of the infrared pole — corresponding to the renormalization in the ultraviolet case — is called mass factorization and works exactly the same way as renormalizing a parameter: in a well–defined scheme we simply subtract the pole from the fixed-order matrix element squared.

From the discussion of the process $qg \to Zq$ we can at least hope that after taking care of all other infrared and ultraviolet divergences the collinear structure of the process $q\bar{q} \to Zg$ will be similar. In this section we will show that we can indeed write all collinear divergences in a universal form, independent of the hard process which we choose as the Drell–Yan process. In the collinear limit, the radiation of additional partons or the splitting into additional partons will be described by universal splitting functions.

Infrared divergences occur for massless particles in the initial or final state, so we need to go through all ways incoming or outgoing gluons and quark can split into each other. The description of the factorized phase space, with which we will start, is common to all these different channels. The first and at the LHC most important case is the splitting of one gluon into two, shown in Figure 2. The two daughter gluons are close to mass shell while the mother has to have a finite positive invariant mass $p_a^2 \gg p_b^2, p_c^2$. We again assign the direction of the momenta as $p_a = -p_b - p_c$, which means we have to take care of minus signs in the particle energies. The kinematics of this approximately collinear process we can describe in terms of the energy fractions z and $1-z$ defined as

$$z = \frac{|E_b|}{|E_a|} = 1 - \frac{|E_c|}{|E_a|} \qquad p_a^2 = (-p_b - p_c)^2 = 2(p_b p_c) = 2z(1-z)(1-\cos\theta)E_a^2 = z(1-z)E_a^2\theta^2 + \mathcal{O}(\theta^4)$$

$$\Leftrightarrow \qquad \theta \equiv \theta_b + \theta_c \simeq \frac{1}{|E_a|}\,\sqrt{\frac{p_a^2}{z(1-z)}}\,, \tag{II.38}$$

in the collinear limit and in terms of the opening angle θ between $\vec{p}_b$ and $\vec{p}_c$. Because $p_a^2 > 0$ we call this final–state splitting configuration time–like branching. For this configuration we can write down the so-called Sudakov decomposition

of the four-momenta

$$-p_a = p_b + p_c = (-zp_a + \beta n + p_T) \; + \; (-(1-z)p_a - \beta n - p_T) \; . \tag{II.39}$$

It defines an arbitrary unit four-vector n, a p_T component orthogonal to the mother momentum p_a and to n, $(p_a p_T) = 0 = (np_T)$, and a free factor β. This way, we can specify n such that it defines the direction of the p_b–p_c decay plane. In this decomposition we can set only one invariant mass to zero, for example that of a radiated gluon $p_c^2 = 0$. The second final state will have a finite invariant mass $p_b^2 \neq 0$.

As specific choice for the three reference four-vectors is

$$p_a = \begin{pmatrix} |E_a| \\ 0 \\ 0 \\ p_{a,3} \end{pmatrix} = |E_a| \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 + \mathcal{O}(\theta) \end{pmatrix} \qquad n = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} \qquad p_T = \begin{pmatrix} 0 \\ p_{T,1} \\ p_{T,2} \\ 0 \end{pmatrix} \; . \tag{II.40}$$

Relative to $\vec{p}_a$ we can split the opening angle θ for massless partons according to Figure 2

$$\theta = \theta_b + \theta_c \qquad \text{and} \qquad \frac{\theta_b}{\theta_c} = \frac{p_T}{|E_b|} \left(\frac{p_T}{|E_c|} \right)^{-1} = \frac{1-z}{z} \qquad \Leftrightarrow \qquad \theta = \frac{\theta_b}{1-z} = \frac{\theta_c}{z} \; . \tag{II.41}$$

The momentum choice in Eq.(II.40) has the additional feature that $n^2 = 0$, which allows us to extract β from the momentum parameterization shown in Eq.(II.39) and the additional condition that $p_c^2 = 0$

$$\begin{aligned} p_c^2 &= (-(1-z)p_a - \beta n - p_T)^2 \\ &= (1-z)^2 p_a^2 + p_T^2 + 2\beta(1-z)(np_a) \\ &= (1-z)^2 p_a^2 + p_T^2 + 4\beta(1-z)|E_a|(1+\mathcal{O}(\theta)) \stackrel{!}{=} 0 \qquad \Leftrightarrow \qquad \beta \simeq -\frac{p_T^2 + (1-z)^2 p_a^2}{4(1-z)|E_a|} \; . \end{aligned} \tag{II.42}$$

Using this specific phase space parameterization we can divide an $(n+1)$-particle process into an n-particle process and a splitting process of quarks and gluons. First, this requires us to split the $(n+1)$-particle phase space into an n-particle phase space and the collinear splitting. The general $(n+1)$-particle phase space separating off the n-particle contribution

$$\begin{aligned} d\Phi_{n+1} &= \cdots \frac{d^3\vec{p}_b}{2(2\pi)^3|E_b|} \; \frac{d^3\vec{p}_c}{2(2\pi)^3|E_c|} \; = \cdots \frac{d^3\vec{p}_a}{2(2\pi)^3|E_a|} \; \frac{d^3\vec{p}_c}{2(2\pi)^3|E_c|} \frac{|E_a|}{|E_b|} \qquad \text{at fixed } p_a \\ &= d\Phi_n \; \frac{dp_{c,3} dp_T p_T d\phi}{2(2\pi)^3|E_c|} \; \frac{1}{z} \\ &= d\Phi_n \; \frac{dp_{c,3} dp_T^2 d\phi}{4(2\pi)^3|E_c|} \; \frac{1}{z} \end{aligned} \tag{II.43}$$

is best expressed in terms of the energy fraction z and the azimuthal angle ϕ. In other words, separating the $(n+1)$-particle space into an n-particle phase space and a $(1 \to 2)$ splitting phase space is possible without any approximation, and all we have to take care of is the correct prefactors in the new parameterization.

Our next task is to translate the phase space parameters $p_{c,3}$ and p_T^2 appearing in Eq.(II.43) into z and p_a^2. Starting from Eq.(II.39) for $p_{c,3}$ with the third components of p_a and p_T given by Eq.(II.40) we insert β from Eq.(II.42) and

obtain

$$\begin{aligned}
\frac{dp_{c,3}}{dz} &= \frac{d}{dz}\left[-(1-z)|E_a|(1+\mathcal{O}(\theta))+\beta\right] = \frac{d}{dz}\left[-(1-z)|E_a|(1+\mathcal{O}(\theta)) - \frac{p_T^2+(1-z)^2p_a^2}{4(1-z)|E_a|}\right] \\
&= |E_a|(1+\mathcal{O}(\theta)) - \frac{p_T^2}{4(1-z)^2E_a} + \frac{p_a^2}{4|E_a|} \\
&= \frac{|E_c|}{1-z}(1+\mathcal{O}(\theta)) - \frac{\theta^2 z^2 E_c^2}{4(1-z)^2E_a} + \frac{z(1-z)E_a^2\theta^2+\mathcal{O}(\theta^4)}{4|E_a|} \qquad \text{using Eq.(II.38) and Eq.(II.41)} \\
&= \frac{|E_c|}{1-z} + \mathcal{O}(\theta) \qquad \Leftrightarrow \qquad \frac{dp_{c,3}}{|E_c|} \simeq \frac{dz}{1-z}\,. \qquad \text{(II.44)}
\end{aligned}$$

In addition to substituting $dp_{c,3}$ by dz in Eq.(II.43) we also replace dp_T^2 with dp_a^2 according to

$$\frac{p_T^2}{p_a^2} = \frac{E_b^2\theta_b^2}{z(1-z)E_a^2\theta^2} = \frac{z^2E_a^2(1-z)^2\theta^2}{z(1-z)E_a^2\theta^2} = z(1-z) \quad \Leftrightarrow \quad dp_T^2 = z(1-z)dp_a^2\,. \qquad \text{(II.45)}$$

This gives us the final result for the separated collinear phase space

$$\boxed{d\Phi_{n+1} = d\Phi_n\,\frac{dz\,dp_a^2\,d\phi}{4(2\pi)^3} = d\Phi_n\,\frac{dz\,dp_a^2}{4(2\pi)^2}}\,, \qquad \text{(II.46)}$$

where in the second step we assume an azimuthal symmetry.

Adding the matrix element to this factorization of the phase space and ignoring the initial–state flux factor which is common to both processes we can now postulate a full factorization for one collinear emission in the collinear limit

$$\begin{aligned}
d\sigma_{n+1} &= \overline{|\mathcal{M}_{n+1}|^2}\; d\Phi_{n+1} \\
&= \overline{|\mathcal{M}_{n+1}|^2}\; d\Phi_n \frac{dp_a^2\,dz}{4(2\pi)^2}\;\; (1+\mathcal{O}(\theta)) \\
&\simeq \frac{2g_s^2}{p_a^2}\;\hat{P}(z)\;\overline{|\mathcal{M}_n|^2}\; d\Phi_n \frac{dp_a^2\,dz}{16\pi^2} \qquad \text{assuming} \quad \overline{|\mathcal{M}_{n+1}|^2} \simeq \frac{2g_s^2}{p_a^2}\;\hat{P}(z)\;\overline{|\mathcal{M}_n|^2}\,. \qquad \text{(II.47)}
\end{aligned}$$

Using $d\sigma_n \sim \overline{|\mathcal{M}_n|^2}\; d\Phi_n$ and $g_s^2 = 4\pi\alpha_s$ we can write this relation in its most common form

$$\boxed{\sigma_{n+1} \simeq \int \sigma_n\;\frac{dp_a^2}{p_a^2}dz\;\frac{\alpha_s}{2\pi}\;\hat{P}(z)}\,. \qquad \text{(II.48)}$$

We can show the assumption of factorizing matrix elements step by step, constructing the appropriate splitting kernels $\hat{P}(z)$ for all different quark and gluon configurations. If Eq.(II.48) really holds true this means that we can compute the $(n+1)$ particle amplitude squared from the n-particle case convoluted with appropriate universal splitting kernels.

As the first parton splitting in QCD we study a gluon splitting into two gluons

$$g(p_a) \to g(p_b) + g(p_c)\,, \qquad \text{(II.49)}$$

also shown in Figure 2. We can compute it in the collinear configuration given in Eq.(II.38) and find

$$\begin{aligned}
\overline{|\mathcal{M}_{n+1}|^2} &= \frac{2g_s^2}{p_a^2}\;\frac{N_c}{2}\;2\left[\frac{z}{1-z} + z(1-z) + \frac{1-z}{z}\right]\;\overline{|\mathcal{M}_n|^2} \\
&\equiv \frac{2g_s^2}{p_a^2}\;\hat{P}_{g\leftarrow g}(z)\;\overline{|\mathcal{M}_n|^2} \qquad \Leftrightarrow \qquad \hat{P}_{g\leftarrow g}(z) = C_A\left[\frac{z}{1-z} + \frac{1-z}{z} + z(1-z)\right]\,, \qquad \text{(II.50)}
\end{aligned}$$

using $C_A = N_c$. The form of the splitting kernel is symmetric when we exchange the two gluons z and $(1-z)$. It diverges if either of the gluons become soft. The notation $\hat{P}_{i\leftarrow j} \sim \hat{P}_{ij}$ is inspired by a matrix notation which we can use to multiply the splitting matrix from the right with the incoming parton vector to get the final parton vector.

Following the logic described above, with this calculation we prove that the factorized form of the $(n+1)$-particle matrix element squared in Eq.(II.47) holds in a Yang-Mills theory.

The same kind of splitting kernel we can compute for the splitting of a gluon into two quarks

$$g(p_a) \to q(p_b) + \bar{q}(p_c) \; . \tag{II.51}$$

In complete analogy to the gluon splitting into two gluons, we can factorize the $(n+1)$-particle matrix element into

$$\begin{aligned}
\overline{|\mathcal{M}_{n+1}|^2} &= \frac{g_s^2}{p_a^2} \, T_R \frac{N_c^2-1}{N_c^2-1} \, \left[(1-2z)^2+1\right] \, \overline{|\mathcal{M}_n|^2} \qquad\qquad \text{with } \mathrm{tr} T^a T^b = T_R \delta^{ab} \text{ and } N_a = 2 \\
&= \frac{2g_s^2}{p_a^2} \, T_R \, \left[z^2+(1-z)^2\right] \, \overline{|\mathcal{M}_n|^2} \\
&\equiv \frac{2g_s^2}{p_a^2} \hat{P}_{q\leftarrow g}(z) \, \overline{|\mathcal{M}_n|^2} \qquad \Leftrightarrow \qquad \hat{P}_{q\leftarrow g}(z) = T_R \left[z^2+(1-z)^2\right] \; ,
\end{aligned} \tag{II.52}$$

with $T_R = 1/2$. This splitting kernel is again symmetric in z and $(1-z)$ because QCD does not distinguish between the outgoing quark and the outgoing antiquark.

The third splitting we compute is the splitting of a quark into a quark and a gluon

$$q(p_a) \to q(p_b) + g(p_c) \; . \tag{II.53}$$

It involves the same quark–quark–gluon vertex as the gluon splitting into two quarks. The factorized matrix element for this channel has the same form as Eq.(II.52), except for the color averaging factor of the now incoming quark,

$$\begin{aligned}
\overline{|\mathcal{M}_{n+1}|^2} &= \frac{g_s^2}{p_a^2} \, \frac{N_c^2-1}{2N_c} \, \frac{(1+z)^2+(1-z)^2}{1-z} \, \overline{|\mathcal{M}_n|^2} \\
&= \frac{2g_s^2}{p_a^2} \, C_F \, \frac{1+z^2}{1-z} \, \overline{|\mathcal{M}_n|^2} \\
&\equiv \frac{2g_s^2}{p_a^2} \hat{P}_{q\leftarrow g}(z) \, \overline{|\mathcal{M}_n|^2} \qquad \Leftrightarrow \qquad \hat{P}_{q\leftarrow q}(z) = C_F \frac{1+z^2}{1-z} \; .
\end{aligned} \tag{II.54}$$

The color factor for gluon radiation off a quark is $C_F = (N^2-1)/(2N)$. The averaging factor $1/N_a = 2$ now is the number of quark spins in the intermediate state. Just switching $z \leftrightarrow (1-z)$ we can finally read off the kernel for a quark splitting written in terms of the final–state gluon

$$\hat{P}_{g\leftarrow q}(z) = C_F \frac{1+(1-z)^2}{z} \; . \tag{II.55}$$

This result finalizes our calculation of all QCD splitting kernels $\hat{P}_{i\leftarrow j}(z)$ between quarks and gluons . As alluded to earlier, similar to ultraviolet divergences which get removed by counter terms these splitting kernels are universal. They do not depend on the hard n-particle matrix element which is part of the original $(n+1)$-particle process. Based on our four results in Eqs.(II.50), (II.52), (II.54), and (II.55) we have by construction of the kernels $\hat{P}$ shown that the collinear factorization Eq.(II.48) holds at this level in perturbation theory.

Before using this splitting property to describe QCD effects at the LHC we need to look at the splitting of partons in the initial state, meaning $|p_a^2|, p_c^2 \ll |p_b^2|$ where p_b is the momentum entering the hard interaction. The difference to the final–state splitting is that now we can consider the split parton momentum $p_b = p_a - p_c$ as a t-channel diagram, so we already know $p_b^2 = t < 0$ from our usual Mandelstam variables argument. This space–like splitting version of

Eq.(II.39) for p_b^2 gives us

$$\begin{aligned} t \equiv p_b^2 &= (-zp_a + \beta n + p_T)^2 \\ &= p_T^2 - 2z\beta(p_a n) && \text{with } p_a^2 = n^2 = (p_a p_T) = (n p_T) = 0 \\ &= p_T^2 + \frac{p_T^2 z}{1-z} && \text{using Eq.(II.42)} \\ &= \frac{p_T^2}{1-z} = -\frac{p_{T,1}^2 + p_{T,2}^2}{1-z} < 0 \; . \end{aligned} \tag{II.56}$$

The calculation of the splitting kernels and matrix elements is the same as for the time–like case, with the one exception that for splitting in the initial state the flow factor has to be evaluated at the reduced partonic energy $E_b = zE_a$ and that the energy fraction entering the parton density needs to be replaced by $x_b \to zx_b$. The factorized matrix element for initial–state splitting then reads just like Eq.(II.48)

$$\sigma_{n+1} = \int \sigma_n \; \frac{dt}{t} dz \; \frac{\alpha_s}{2\pi} \; \hat{P}(z) \; . \tag{II.57}$$

How to use this property to make statements about the quark and gluon content in the proton will be the focus of the next section.

D. DGLAP equation

Before we include the quantum effects from parton splitting in LHC computations, let us briefly review the classical, statisitical picture of parton densities. At hadron colliders the energy distribution of incoming quarks as parts of the colliding protons has to be taken into account. We first assume that quarks move collinearly with the surrounding proton such that at the LHC incoming partons have zero p_T. Under that condition we can define a probability distribution for finding a parton just depending on the respective fraction of the proton's momentum. For this momentum fraction $x = 0 \cdots 1$ the parton density function (pdf) is written as $f_i(x)$, where i denotes the different partons in the proton, for our purposes u, d, c, s, g and, depending on the details, b. All incoming partons we assume to be massless.

In contrast to so-called structure functions a pdf is not an observable. It is a distribution in the mathematical sense, which means it has to produce reasonable results when we integrate it together with a test function. Different parton densities have very different behavior — for the valence quarks (uud) they peak somewhere around $x \lesssim 1/3$, while the gluon pdf is small at $x \sim 1$ and grows very rapidly towards small x. For some typical part of the relevant parameter space ($x = 10^{-3} \cdots 10^{-1}$) the gluon density roughly scales like $f_g(x) \propto x^{-2}$. Towards smaller x values it becomes even steeper. This steep gluon distribution was initially not expected and means that for small enough x LHC processes will dominantly be gluon fusion processes.

While we cannot actually compute parton distribution functions $f_i(x)$ as a function of the momentum fraction x there are a few predictions we can make based on symmetries and properties of the hadrons. Such arguments for example lead to sum rules:

The parton distributions inside an antiproton are linked to those inside a proton through the CP symmetry, which is an exact symmetry of QCD. Therefore, we know that

$$f_q^{\bar{p}}(x) = f_{\bar{q}}(x) \qquad\qquad f_{\bar{q}}^{\bar{p}}(x) = f_q(x) \qquad\qquad f_g^{\bar{p}}(x) = f_g(x) \tag{II.58}$$

for all values of x. If the proton consists of three valence quarks uud, plus quantum fluctuations from the vacuum which can either involve gluons or quark–antiquark pairs, the contribution from the sea quarks has to be symmetric in quarks and antiquarks. The expectation values for the signed numbers of up and down quarks inside a proton have to fulfill

$$\langle N_u \rangle = \int_0^1 dx \; (f_u(x) - f_{\bar{u}}(x)) = 2 \qquad\qquad \langle N_d \rangle = \int_0^1 dx \; (f_d(x) - f_{\bar{d}}(x)) = 1 \; . \tag{II.59}$$

Similarly, the total momentum of the proton has to consist of sum of all parton momenta. We can write this as the

expectation value of $\sum x_i$

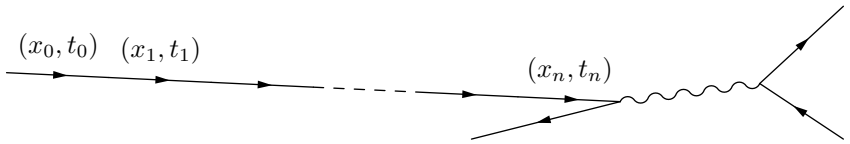
$$\langle \sum x_i \rangle = \int_0^1 dx\, x \left(\sum_q f_q(x) + \sum_{\bar{q}} f_{\bar{q}}(x) + f_g(x) \right) = 1 \tag{II.60}$$

What makes this prediction interesting is that we can compute the same sum only taking into account the measured quark and antiquark parton densities. We find that this modified momentum sum rule only comes to 1/2. Half of the proton momentum is then carried by gluons.

We can use everything we now know about collinear parton splitting to describe incoming partons at hadron colliders. For example in $pp \to Z$ production incoming partons inside the protons transform into each other via collinear splitting until they enter the Z production process as quarks. Taking Eq.(II.57) seriously, a parton density should depend on two parameters, the final energy fraction and the virtuality,

$$f(x, -t) \; . \tag{II.61}$$

The second parameter t is new compared to the purely probabilistic picture described above. To study the parton density as a function of these two parameters, we start with a quark inside the proton with an energy fraction x_0, as it enters the hadronic phase space integral. Since this quark is confined inside the proton it can only have small transverse momentum, which means its four-momentum squared t_0 is negative and its absolute value $|t_0|$ is small. For incoming partons which on–shell have $p^2 = 0$ it gives the distance to the mass shell. Let us simplify our kinematic argument by assuming that there exists only one type of splitting, namely successive gluon radiation off an incoming quark, where the outgoing gluons are not relevant



In that case each collinear gluon radiation will decrease the quark energy $x_{j+1} < x_j$ and increase its virtuality $|t_{j+1}| = -t_{j+1} > -t_j = |t_j|$ through its recoil.

From the last section we know what the successive splitting means in terms of splitting probabilities. We can describe how the parton density $f(x, -t)$ evolves in the $(x - t)$ plane as depicted in Figure 3. The starting point (x_0, t_0) is at least probabilistically given by the energy and kind of the hadron, for example the proton. For a given small virtuality $|t_0|$ we start at some kind of fixed x_0 distribution. We then interpret each branching as a step strictly downward in $x_j \to x_{j+1}$ where the t value we assign to this step is the ever increasing virtuality $|t_{j+1}|$ after the branching. Each splitting means a synchronous shift in x and t, so the actual path in the $(x - t)$ plane really consists of discrete points. The probability of such a splitting to occur is given by $\hat{P}_{q \leftarrow q}(z) \equiv \hat{P}(z)$ as it appears in Eq.(II.57)

$$\frac{\alpha_s}{2\pi} \hat{P}(z) \frac{dt}{t} dz \; . \tag{II.62}$$

In this picture we consider this probability a smooth function in t and z. At the end of the path we will probe this evolved parton density, where x_n and t_n enter the hard scattering process and its energy–momentum conservation.

When we convert a partonic into a hadronic cross section numerically we need to specify the probability of the parton density $f(x, -t)$ residing in an infinitesimal square $[x_j, x_j + \delta x]$ and, if this second parameter has anything to do with physics, $[|t_j|, |t_j| + \delta t]$. Using our (x, t) plane we compute the flows into this square and out of this square, which together define the net shift in f in the sense of a differential equation, similar to the derivation of Gauss' theorem for vector fields inside a surface

$$\delta f_{\text{in}} - \delta f_{\text{out}} = \delta f(x, -t) \; . \tag{II.63}$$

We compute the incoming and outgoing flows from the history of the (x, t) evolution. At this stage our picture becomes a little subtle; the way we define the path between two splittings in Figure 3 it can enter and leave the square either vertically or horizontally. Because we do not consider the movement in the (x, t) plane continuous we can choose this direction as vertical or horizontal. Because we want to arrive at a differential equation in t we choose the vertical drop, such that the area the incoming and outgoing flows see is given by δt. If we define a splitting as

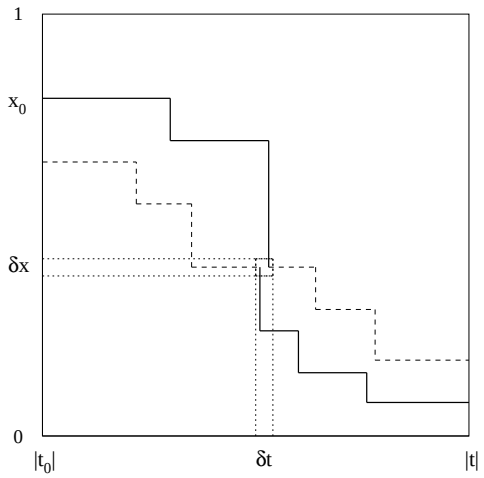


FIG. 3: Path of an incoming parton in the $(x-t)$ plane. Because we define t as a negative number its axis is labelled $|t|$.

such a vertical drop in x at the target value t_{j+1} an incoming path hitting the square at some value t can come from any x value above the square. Using this convention and following the fat solid lines in Figure 3 the vertical flow into the square (x,t) square is proportional to δt and has the form

$$\begin{aligned}
\delta f_{\rm in}(-t) &= \frac{\delta t}{t}\int_x^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,\hat{P}(z) f\left(\frac{x}{z},-t\right) \\
&= \frac{\delta t}{t}\int_0^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,\hat{P}(z) f\left(\frac{x}{z},-t\right) \qquad \text{assuming } f(x',-t)=0 \text{ for } x'>1 \\
&\equiv \delta t\,\left(\frac{\alpha_s \hat{P}}{2\pi t}\otimes f\right)(x,-t)\,. \qquad \text{(II.64)}
\end{aligned}$$

In the last step we use the definition of a convolution

$$(f\otimes g)(x) = \int_0^1 dx_1 dx_2 f(x_1) g(x_2)\,\delta(x-x_1x_2) = \int_0^1 \frac{dx_1}{x_1} f(x_1) g\left(\frac{x}{x_1}\right) = \int_0^1 \frac{dx_2}{x_2} f\left(\frac{x}{x_2}\right) g(x_2)\,. \qquad \text{(II.65)}$$

The outgoing flow we define in complete analogy, again leaving the infinitesimal square vertically. Following the fat solid line in Figure 3 it is also proportional to δt

$$\delta f_{\rm out}(-t) = \delta t\int_0^1 dy \frac{\alpha_s \hat{P}(y)}{2\pi t}\,f(x,-t) = \frac{\delta t}{t} f(x,-t)\int_0^1 dy\,\frac{\alpha_s}{2\pi}\,\hat{P}(y)\,. \qquad \text{(II.66)}$$

The corresponding y integration, unlike the z integration for the incoming flow is not a convolution. This integration appears because we do not know the normalization of $\hat{P}(z)$ distribution which we interpret as a probability. The reason why it is not a convolution is that for the outgoing flow we know the starting condition and integrate over the final configurations; this aspect will become important later. Combining Eq.(II.64) and Eq.(II.66) we can compute the net change in the parton density of the quarks as

$$\begin{aligned}
\delta f(x,-t) &= \frac{\delta t}{t}\left[\int_0^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,\hat{P}(z)\,f\left(\frac{x}{z},-t\right) - \int_0^1 dy\,\frac{\alpha_s}{2\pi}\,\hat{P}(y)\,f(x,-t)\right] \\
&= \frac{\delta t}{t}\int_0^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\left[\hat{P}(z) - \delta(1-z)\int_0^1 dy \hat{P}(y)\right] f\left(\frac{x}{z},-t\right) \equiv \frac{\delta t}{t}\int_x^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,\hat{P}(z)_+\,f\left(\frac{x}{z},-t\right) \\
\Leftrightarrow \qquad \frac{\delta f(x,-t)}{\delta(-t)} &= \frac{1}{(-t)}\int_x^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,\hat{P}(z)_+\,f\left(\frac{x}{z},-t\right)\,, \qquad \text{(II.67)}
\end{aligned}$$

again assuming $f(x) = 0$ for $x > 1$, strictly speaking requiring α_s to only depend on t but not on z, and using the specifically defined plus subtraction

$$\boxed{F(z)_+ \equiv F(z) - \delta(1-z) \int_0^1 dy\, F(y)} \qquad \text{or} \qquad \int_0^1 dz\, \frac{f(z)}{(1-z)_+} = \int_0^1 dz\, \left(\frac{f(z)}{1-z} - \frac{f(1)}{1-z} \right) . \tag{II.68}$$

For the second form above we choose $F(z) = 1/(1-z)$, multiply it with an arbitrary test function $f(z)$ and integrate over z. In contrast to the original z integral the plus–subtracted integral is by definition finite in the limit $z \to 1$, where some of the splitting kernels diverge. For example, the quark splitting kernel including the plus prescription becomes $C_F((1+z^2)/(1-z))_+$. At this stage the plus prescription is simply a convenient way of writing a complicated combination of splitting kernels, but we will see that it also has a physics meaning.

Next, we check that the plus prescription indeed acts as a regularization technique for the parton densities. Obviously, the integral over $f(z)/(1-z)$ is divergent at the boundary $z \to 1$, which we know we can cure using dimensional regularization. The special case $f(z) = 1$ illustrates how dimensional regularization of infrared divergences in the phase space integration Eq.(II.37) works

$$\int_0^1 dz\, \frac{1}{(1-z)^{1-\epsilon}} = \int_0^1 dz\, \frac{1}{z^{1-\epsilon}} = \frac{z^\epsilon}{\epsilon}\bigg|_0^1 = \frac{1}{\epsilon} \qquad \text{with } \epsilon > 0 , \tag{II.69}$$

for $4 + 2\epsilon$ dimensions. This change in sign avoids the analytic continuation of the usual value $n = 4 - 2\epsilon$ to $\epsilon < 0$. The dimensionally regularized integral in analogy to Eq.(II.68) we can write as

$$\begin{aligned}
\int_0^1 dz\, \frac{f(z)}{(1-z)^{1-\epsilon}} &= \int_0^1 dz\, \frac{f(z)-f(1)}{(1-z)^{1-\epsilon}} + f(1) \int_0^1 dz\, \frac{1}{(1-z)^{1-\epsilon}} \\
&= \int_0^1 dz\, \frac{f(z)-f(1)}{1-z} \left(1 + \mathcal{O}(\epsilon)\right) + \frac{f(1)}{\epsilon} \\
&= \int_0^1 dz\, \frac{f(z)}{(1-z)_+} \left(1 + \mathcal{O}(\epsilon)\right) + \frac{f(1)}{\epsilon} \qquad \text{by definition} \\
\Leftrightarrow \qquad \int_0^1 dz\, \frac{f(z)}{(1-z)^{1-\epsilon}} - \frac{f(1)}{\epsilon} &= \int_0^1 dz\, \frac{f(z)}{(1-z)_+} \left(1 + \mathcal{O}(\epsilon)\right) .
\end{aligned} \tag{II.70}$$

The dimensionally regularized integral minus the pole, *i.e.* the finite part of the dimensionally regularized integral, is the same as the plus–subtracted integral modulo terms of the order ϵ. The third line in Eq.(II.70) shows that the difference between a dimensionally regularized splitting kernel and a plus–subtracted splitting kernel manifests itself as terms proportional to $\delta(1-z)$. Physically, they represent contributions to a soft–radiation phase space integral.

Before we move on introducing a gluon density we can slightly reformulate the splitting kernel $\hat{P}_{q \leftarrow q}$ in Eq.(II.54). If the plus prescription regularizes the pole at $z \to 1$, what happens when we include the numerator of the regularized function, *e.g.* the quark splitting kernel? The finite difference between these results is

$$\begin{aligned}
\left(\frac{1+z^2}{1-z} \right)_+ - (1+z^2) \left(\frac{1}{1-z} \right)_+ &= \frac{1+z^2}{1-z} - \delta(1-z) \int_0^1 dy\, \frac{1+y^2}{1-y} - \frac{1+z^2}{1-z} + \delta(1-z) \int_0^1 dy\, \frac{1+z^2}{1-y} \\
&= -\delta(1-z) \int_0^1 dy\, \left(\frac{1+y^2}{1-y} - \frac{2}{1-y} \right) \\
&= \delta(1-z) \int_0^1 dy\, \frac{y^2-1}{y-1} = \delta(1-z) \int_0^1 dy\, (y+1) = \frac{3}{2} \delta(1-z) .
\end{aligned} \tag{II.71}$$

We can therefore write the quark's splitting kernel in two equivalent ways

$$\boxed{P_{q \leftarrow q}(z) = C_F \left(\frac{1+z^2}{1-z} \right)_+ = C_F \left[\frac{1+z^2}{(1-z)_+} + \frac{3}{2} \delta(1-z) \right]} . \tag{II.72}$$

Going back to our evolution in (x,t) space, the infinitesimal version of Eq.(II.67) is the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi or DGLAP integro-differential equation which describes the scale dependence of the quark parton density. As we already know quarks do not only appear in $q \to q$ splitting, but also in gluon splitting. Therefore, we generalize Eq.(II.67) to include the full set of QCD partons, *i.e.* quarks and gluons. This generalization involves a sum over all allowed splittings and the plus–subtracted splitting kernels. For the quark density on the left hand side it is

$$\boxed{\frac{df_q(x,-t)}{d\log(-t)} = -t\,\frac{df_q(x,-t)}{d(-t)} = \sum_{j=q,g}\int_x^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,P_{q\leftarrow j}(z)\,f_j\left(\frac{x}{z},-t\right)} \qquad \text{with} \quad P_{q\leftarrow j}(z) \equiv \hat{P}_{q\leftarrow j}(z)_+ \,. \tag{II.73}$$

Expanding Eq.(II.67) beyond just gluon radiation off hard quarks, we add all relevant parton indices and splittings which lead to a quark density. Including the splitting from gluon to quark we arrive at

$$\begin{aligned}\delta f_q(x,-t) = \frac{\delta t}{t}\,\Bigg[&\int_0^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,\hat{P}_{q\leftarrow q}(z)\,f_q\left(\frac{x}{z},-t\right) + \int_0^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,\hat{P}_{q\leftarrow g}(z)\,f_g\left(\frac{x}{z},-t\right)\\ &- \int_0^1 dy\,\frac{\alpha_s}{2\pi}\,\hat{P}_{q\leftarrow q}(y)\,f_q(x,-t)\Bigg]\,.\end{aligned} \tag{II.74}$$

Of the three terms on the right hand side the first and the third together define the plus–subtracted splitting kernel $P_{q\leftarrow q}(z)$, just following the argument above. The second term is a proper convolution and the only term proportional to the gluon parton density. Quarks can be produced in gluon splitting but cannot vanish into it. Therefore, we have to identify the last term in Eq.(II.74) with $P_{q\leftarrow g}$, without adding a plus–regulator

$$\boxed{P_{q\leftarrow g}(z) \equiv \hat{P}_{q\leftarrow g}(z) = T_R\left[z^2+(1-z)^2\right]}\,. \tag{II.75}$$

In principle, the splitting kernel $\hat{P}_{g\leftarrow q}$ also generates a quark, in addition to the final–state gluon. However, comparing this to the terms proportional to $\hat{P}_{q\leftarrow q}$ they both arise from the same splitting, namely a quark density leaving the infinitesimal square in the $(x-t)$ plane via the splitting $q \to qg$. Including the additional $\hat{P}_{g\leftarrow q}(y)$ would be double counting and should not appear, as the notation $g \leftarrow q$ already suggests.

The second QCD parton density we have to study is the gluon density. The incoming contribution to the infinitesimal square is given by the sum of four splitting scenarios each leading to a gluon with virtuality $-t_{j+1}$

$$\begin{aligned}\delta f_{\text{in}}(-t) &= \frac{\delta t}{t}\int_0^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,\left[\hat{P}_{g\leftarrow g}(z)\left(f_g\left(\frac{x}{z},-t\right)+f_g\left(\frac{x}{1-z},-t\right)\right)+\hat{P}_{g\leftarrow q}(z)\left(f_q\left(\frac{x}{z},-t\right)+f_{\bar{q}}\left(\frac{x}{z},-t\right)\right)\right]\\ &= \frac{\delta t}{t}\int_0^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,\left[2\hat{P}_{g\leftarrow g}(z) f_g\left(\frac{x}{z},-t\right)+\hat{P}_{g\leftarrow q}(z)\left(f_q\left(\frac{x}{z},-t\right)+f_{\bar{q}}\left(\frac{x}{z},-t\right)\right)\right]\,,\end{aligned} \tag{II.76}$$

using $P_{g\leftarrow\bar{q}} = P_{g\leftarrow q}$ in the first line and $P_{g\leftarrow g}(1-z) = P_{g\leftarrow g}(z)$ in the second. To leave the volume element in the (x,t) space a gluon can either split into two gluons or radiate one of n_f light-quark flavors. Combining the incoming and outgoing flows we find

$$\begin{aligned}\delta f_g(x,-t) =& \frac{\delta t}{t}\int_0^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,\left[2\hat{P}_{g\leftarrow g}(z) f_g\left(\frac{x}{z},-t\right)+\hat{P}_{g\leftarrow q}(z)\left(f_q\left(\frac{x}{z},-t\right)+f_{\bar{q}}\left(\frac{x}{z},-t\right)\right)\right]\\ &-\frac{\delta t}{t}\int_0^1 dy\,\frac{\alpha_s}{2\pi}\,\left[\hat{P}_{g\leftarrow g}(y)+n_f\hat{P}_{q\leftarrow g}(y)\right] f_g(x,-t)\end{aligned} \tag{II.77}$$

We have to evaluate the four terms in this expression one after the other. Unlike in the quark case they do not immediately correspond to regularizing the diagonal splitting kernel using the plus prescription.

First, there exists a contribution to δf_{in} proportional to f_q or $f_{\bar{q}}$ which is not matched by the outgoing flow. From the quark case we already know how to deal with it. For the corresponding splitting kernel there is no regularization

through the plus prescription needed, so we define

$$\boxed{P_{g\leftarrow q}(z) \equiv \hat{P}_{g\leftarrow q}(z) = C_F \frac{1+(1-z)^2}{z}} \,. \tag{II.78}$$

This ensures that the off-diagonal contribution to the gluon density is taken into account when we extend Eq.(II.73) to a combined quark/antiquark and gluon form. The structure of the DGLAP equation implies that the two off-diagonal splitting kernels do not include any plus prescription $\hat{P}_{i\leftarrow j} = P_{i\leftarrow j}$. We could have expected this, because off-diagonal kernels are finite in the soft limit, $z \to 1$. Applying a plus prescription would only have modified the splitting kernels at the isolated (zero-measure) point $y = 1$ which for a finite value of the integrand does not affect the integral on the right hand side of the DGLAP equation.

Second, the y integral describing the gluon splitting into a quark pair we can compute directly,

$$\begin{aligned}
-\int_0^1 dy\, \frac{\alpha_s}{2\pi}\, n_f\, \hat{P}_{q\leftarrow g}(y) &= -\frac{\alpha_s}{2\pi}\, n_f\, T_R \int_0^1 dy\, \left[1-2y+2y^2\right] \qquad \text{using Eq.(II.75)} \\
&= -\frac{\alpha_s}{2\pi}\, n_f\, T_R \left[y - y^2 + \frac{2y^3}{3}\right]_0^1 \\
&= -\frac{2}{3}\, \frac{\alpha_s}{2\pi}\, n_f\, T_R \,.
\end{aligned} \tag{II.79}$$

Finally, the terms proportional to the purely gluonic splitting $P_{g\leftarrow g}$ appearing in Eq.(II.77) require some more work. The y integral coming from the outgoing flow has to consist of a finite term and a term we can use to define the plus prescription for $\hat{P}_{g\leftarrow g}$. We can compute the integral as

$$\begin{aligned}
-\int_0^1 dy\, \frac{\alpha_s}{2\pi}\, \hat{P}_{g\leftarrow g}(y) &= -\frac{\alpha_s}{2\pi}\, C_A \int_0^1 dy\, \left[\frac{y}{1-y} + \frac{1-y}{y} + y(1-y)\right] \qquad \text{using Eq.(II.50)} \\
&= -\frac{\alpha_s}{2\pi}\, C_A \int_0^1 dy\, \left[\frac{2y}{1-y} + y(1-y)\right] \\
&= -\frac{\alpha_s}{2\pi}\, C_A \int_0^1 dy\, \left[\frac{2(y-1)}{1-y} + y(1-y)\right] - \frac{\alpha_s}{2\pi}\, C_A \int_0^1 dy\, \frac{2}{1-y} \\
&= -\frac{\alpha_s}{2\pi}\, C_A \int_0^1 dy\, \left[-2+y-y^2\right] - \frac{\alpha_s}{2\pi}\, 2C_A \int_0^1 dz\, \frac{1}{1-z} \\
&= -\frac{\alpha_s}{2\pi}\, C_A \left[-2+\frac{1}{2}-\frac{1}{3}\right] - \frac{\alpha_s}{2\pi}\, 2C_A \int_0^1 dz\, \frac{1}{1-z} \\
&= \frac{\alpha_s}{2\pi}\, \frac{11}{6}\, C_A - \frac{\alpha_s}{2\pi}\, 2C_A \int_0^1 dz\, \frac{1}{1-z} \,.
\end{aligned} \tag{II.80}$$

The second term in this result is what we need to replace the first term in the splitting kernel of Eq.(II.50) proportional to $1/(1-z)$ by $1/(1-z)_+$. We can see this using $f(z) = z$ and correspondingly $f(1) = 1$ in Eq.(II.68). The two finite terms in Eq.(II.79) and Eq.(II.80) we have to include in the definition of $\hat{P}_{g\leftarrow g}$ ad hoc. Because the regularized splitting kernel appear inside a convolution the two finite terms require an additional term $\delta(1-z)$. Collecting all of them we arrive at

$$\boxed{P_{g\leftarrow g}(z) = 2C_A \left(\frac{z}{(1-z)_+} + \frac{1-z}{z} + z(1-z)\right) + \frac{11}{6}\, C_A\, \delta(1-z) - \frac{2}{3}\, n_f\, T_R\, \delta(1-z)} \,. \tag{II.81}$$

This result concludes our computation of all four regularized splitting functions which appear in the DGLAP equation Eq.(II.73).

Before discussing and solving the DGLAP equation, let us briefly recapitulate: for the full quark and gluon particle content of QCD we have derived the DGLAP equation which describes a factorization scale dependence of the quark and gluon parton densities. The universality of the splitting kernels is obvious from the way we derive them — no information on the n-particle process ever enters the derivation.

The DGLAP equation is formulated in terms of four splitting kernels of gluons and quarks which are linked to the

splitting probabilities, but which for the DGLAP equation have to be regularized. With the help of a plus–subtraction all kernels $P_{i\leftarrow j}(z)$ become finite, including in the soft limit $z \to 1$. However, splitting kernels are only regularized when needed, so the finite off-diagonal quark–gluon and gluon–quark splittings are unchanged. This means the plus prescription really acts as an infrared renormalization, moving universal infrared divergences into the definition of the parton densities. The original collinear divergence has vanished as well.

The only approximation we make in the computation of the splitting kernels is that in the y integrals we implicitly assume that the running coupling α_s does not depend on the momentum fraction. In its standard form and in terms of the factorization scale $\mu_F^2 \equiv -t$ the DGLAP equation reads

$$\boxed{\frac{df_i(x,\mu_F)}{d\log\mu_F^2} = \sum_j \int_x^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,P_{i\leftarrow j}(z)\,f_j\left(\frac{x}{z},\mu_F\right) = \frac{\alpha_s}{2\pi}\,\sum_j \left(P_{i\leftarrow j}\otimes f_j\right)(x,\mu_F)}\,. \tag{II.82}$$

E. Collinear logarithms

Solving the integro-differential DGLAP equation Eq.(II.82) for the parton densities is clearly beyond the scope of this writeup. Nevertheless, we will sketch how we would approach this. This will give us some information on the structure of its solutions which we need to understand the physics of the DGLAP equation.

One simplification we can make is to postulate eigenvalues in parton space and solve the equation for them. This gets rid of the sum over partons on the right hand side. One such parton density is the non–singlet parton density, defined as the difference

$$f_q^{\rm NS} = f_u - f_{\bar u}\,. \tag{II.83}$$

Since gluons cannot distinguish between massless quarks and antiquarks, the gluon contribution to their evolution cancels. This will be true at arbitrary loop order, since flavor symmetries commute with the QCD gauge group. The corresponding DGLAP equation with leading order splitting kernels now reads

$$\frac{df_q^{\rm NS}(x,\mu_F)}{d\log\mu_F^2} = \int_x^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,P_{q\leftarrow q}(z)\,f_q^{\rm NS}\left(\frac{x}{z},\mu_F\right) = \frac{\alpha_s}{2\pi}\,\left(P_{q\leftarrow q}\otimes f_q^{\rm NS}\right)(x,\mu_F)\,. \tag{II.84}$$

To solve it we need a transformation which simplifies a convolution, leading us to the Mellin transform. Starting from a function $f(x)$ of a real variable x we define the Mellin transform into moment space m

$$\boxed{\mathcal{M}[f](m) \equiv \int_0^1 dx\, x^{m-1} f(x)} \qquad\qquad f(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c-i\infty} dm\,\frac{\mathcal{M}[f](m)}{x^m}\,, \tag{II.85}$$

where for the back transformation we choose an arbitrary appropriate constant $c > 0$, such that the integration contour for the inverse transformation lies to the right of all singularities of the analytic continuation of $\mathcal{M}[f](m)$. The Mellin transform of a convolution is the product of the two Mellin transforms, which gives us the transformed DGLAP equation

$$\begin{aligned}
&\mathcal{M}[P_{q\leftarrow q}\otimes f_q^{\rm NS}](m) = \mathcal{M}\left[\int_0^1 \frac{dz}{z} P_{q\leftarrow q}\left(\frac{x}{z}\right) f_q^{\rm NS}(z)\right](m) = \mathcal{M}[P_{q\leftarrow q}](m)\;\mathcal{M}[f_q^{\rm NS}](m,\mu_F)\\
\Rightarrow\qquad &\frac{d\mathcal{M}[f_q^{\rm NS}](m,\mu_F)}{d\log\mu_F^2} = \frac{\alpha_s}{2\pi}\,\mathcal{M}[P_{q\leftarrow q}](m)\quad \mathcal{M}[f_q^{\rm NS}](m,\mu_F)\,.
\end{aligned} \tag{II.86}$$

This simple linear, first-order differential equation has the solution

$$\begin{aligned}\mathcal{M}[f_q^{\rm NS}](m,\mu_F) &= \mathcal{M}[f_q^{\rm NS}](m,\mu_{F,0})\ \exp\left(\frac{\alpha_s}{2\pi}\,\mathcal{M}[P_{q\leftarrow q}](m)\log\frac{\mu_F^2}{\mu_{F,0}^2}\right)\\ &= \mathcal{M}[f_q^{\rm NS}](m,\mu_{F,0})\ \left(\frac{\mu_F^2}{\mu_{F,0}^2}\right)^{\frac{\alpha_s}{2\pi}\mathcal{M}[P_{q\leftarrow q}](m)}\\ &\equiv \mathcal{M}[f_q^{\rm NS}](m,\mu_{F,0})\ \left(\frac{\mu_F^2}{\mu_{F,0}^2}\right)^{\frac{\alpha_s}{2\pi}\gamma(m)}\quad,\end{aligned} \tag{II.87}$$

defining $\gamma(m) = \mathcal{M}[P](m)$.

The solution given by Eq.(II.87) still has the complication that it includes μ_F and α_s as two free parameters. To simplify our result we can include $\alpha_s(\mu_R^2)$ in the running of the DGLAP equation and identify the renormalization scale μ_R of the strong coupling with the factorization scale $\mu_F = \mu_R \equiv \mu$. This allows us to replace $\log\mu^2$ in the DGLAP equation by α_s, including the leading order Jacobian. This is clearly correct for all one-scale problems where we have no freedom to choose either of the two scales. We find

$$\frac{d}{d\log\mu^2} = \frac{d\log\alpha_s}{d\log\mu^2}\,\frac{d}{d\log\alpha_s} = \frac{1}{\alpha_s}\,\frac{d\alpha_s}{d\log\mu^2}\,\frac{d}{d\log\alpha_s} = -\alpha_s b_0\,\frac{d}{d\log\alpha_s}\,. \tag{II.88}$$

This additional factor of α_s on the left hand side will cancel the factor α_s on the right hand side of the DGLAP equation Eq.(II.86)

$$\begin{aligned}\frac{d\mathcal{M}[f_q^{\rm NS}](m,\mu)}{d\log\alpha_s} &= -\frac{1}{2\pi b_0}\gamma(m)\ \mathcal{M}[f_q^{\rm NS}](m,\mu)\\ \Rightarrow\qquad \mathcal{M}[f_q^{\rm NS}](m,\mu) &= \mathcal{M}[f_q^{\rm NS}](m,\mu_0)\ \exp\left(-\frac{1}{2\pi b_0}\,\gamma(m)\log\frac{\alpha_s(\mu^2)}{\alpha_s(\mu_0^2)}\right)\\ &= \mathcal{M}[f_q^{\rm NS}](m,\mu_{F,0})\ \left(\frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)}\right)^{\frac{\gamma(m)}{2\pi b_0}}\quad.\end{aligned} \tag{II.89}$$

Among other things, in this derivation we neglect that some splitting functions have singularities and therefore the Mellin transform is not obviously well defined. Our convolution is not really a convolution either, because we cut it off at Q_0^2 etc; but the final structure in Eq.(II.89) really holds.

Because we will need it in the next section we emphasize that the same kind of solution appears in pure Yang–Mills theory, *i.e.* in QCD without quarks. Looking at the different color factors in QCD this limit can also be derived as the leading terms in N_c. In that case there also exists only one splitting kernel defining an anomalous dimension γ. We find in complete analogy to Eq.(II.89)

$$\boxed{\mathcal{M}[f_g](m,\mu) = \mathcal{M}[f_g](m,\mu_0)\ \left(\frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)}\right)^{\frac{\gamma(m)}{2\pi b_0}}}\,. \tag{II.90}$$

To remind ourselves that in this derivation we unify the renormalization and factorization scales we denote them just as μ. This solution to the DGLAP equation is not completely determined: as a solution to a differential equation it also includes an integration constant which we express in terms of μ_0. The DGLAP equation therefore does not determine parton densities, it only describes their evolution from one scale μ_F to another, just like a renormalization group equation in the ultraviolet.

Remembering how we arrive at the DGLAP equation we notice an analogy to the case of ultraviolet divergences and the running coupling. We start from universal infrared divergences. We describe them in terms of splitting functions which we regularize using the plus prescription. The DGLAP equation plays the role of a renormalization group equation for example for the running coupling. It links parton densities evaluated at different scales μ_F.

In analogy to the scaling logarithms considered in Section II B we now test if we can point to a type of logarithm the DGLAP equation resums by reorganizing our perturbative series of parton splitting. To identify these resummed logarithms we build a physical model based on collinear splitting, but without using the DGLAP equation. We

start from the basic equation defining the physical picture of parton splitting in Eq.(II.48). Only taking into account gluons in pure Yang–Mills theory it precisely corresponds to the starting point of our discussion leading to the DGLAP equation, schematically written as

$$\sigma_{n+1} = \int \sigma_n \, \frac{dt}{t} dz \, \frac{\alpha_s}{2\pi} \, \hat{P}_{g\leftarrow g}(z) \, . \tag{II.91}$$

This form of collinear factorization does not include parton densities and only applies to final state splittings. To include initial state splittings we need a definition of the virtuality variable t. If we remember that $t = p_b^2 < 0$ we can follow Eq.(II.56) and introduce a positive transverse momentum variable $\vec{p}_T^{\,2}$ in the usual Sudakov decomposition, such that

$$-t = -\frac{p_T^2}{1-z} = \frac{\vec{p}_T^{\,2}}{1-z} > 0 \qquad\qquad \Rightarrow \quad \frac{dt}{t} = \frac{dp_T^2}{p_T^2} = \frac{d\vec{p}_T^{\,2}}{\vec{p}_T^{\,2}} \, . \tag{II.92}$$

From the definition of p_T in Eq.(II.39) we see that $\vec{p}_T^{\,2}$ is really the transverse three-momentum of of the parton pair after splitting.

Extending the single parton radiation we consider a ladder of successive splittings of one gluon into two. For a moment, we forget about the actual parton densities and assume that they are part of the hadronic cross section σ_n. In the collinear limit the appropriate convolution gives us

$$\sigma_{n+1}(x,\mu_F) = \int_{x_0}^1 \frac{dx_n}{x_n} \, \hat{P}_{g\leftarrow g}\left(\frac{x}{x_n}\right) \sigma_n(x_n,\mu_0) \int_{\mu_0^2}^{\mu_F^2} \frac{d\vec{p}_{T,n}^{\,2}}{\vec{p}_{T,n}^{\,2}} \frac{\alpha_s(\mu_R^2)}{2\pi} \, . \tag{II.93}$$

The dz in Eq.(II.91) we replace by the proper convolution $\hat{P} \otimes \sigma_n$, evaluated at the momentum fraction x. Because the splitting kernel is infrared divergent we cut off the convolution integral at x_0. Similarly, the transverse momentum integral is bounded by an infrared cutoff μ_0 and the physical external scale μ_F. This is the range in which an additional collinear radiation is included in σ_{n+1}.

For splitting the two integrals in Eq.(II.93) it is crucial that μ_0 is the only scale the matrix element σ_n depends on. The other integration variable, the transverse momentum, does not feature in σ_n because collinear factorization is defined in the limit $\vec{p}_T^{\,2} \to 0$. For α_s we will see in the next step how μ_R can depend on the transverse momentum. All through the argument of this subsection we should keep in mind that we are looking for assumptions which allow us to solve Eq.(II.93) and compare the result to the solution of the DGLAP equation. In other words, these assumptions we will turn into a physics picture of the DGLAP equation and its solutions.

Making μ_F the global upper boundary of the transverse momentum integration for collinear splitting is our first assumption. We can then apply the recursion formula in Eq.(II.93) iteratively

$$\begin{aligned} \sigma_{n+1}(x,\mu_F) \sim & \int_{x_0}^1 \frac{dx_n}{x_n} \, \hat{P}_{g\leftarrow g}\left(\frac{x}{x_n}\right) \cdots \int_{x_0}^1 \frac{dx_1}{x_1} \, \hat{P}_{g\leftarrow g}\left(\frac{x_2}{x_1}\right) \, \sigma_1(x_1,\mu_0) \\ & \times \int_{\mu_0}^{\mu_F} \frac{d\vec{p}_{T,n}^{\,2}}{\vec{p}_{T,n}^{\,2}} \frac{\alpha_s(\mu_R^2)}{2\pi} \cdots \int_{\mu_0} \frac{d\vec{p}_{T,1}^{\,2}}{\vec{p}_{T,1}^{\,2}} \frac{\alpha_s(\mu_R^2)}{2\pi} \, . \end{aligned} \tag{II.94}$$

The two sets of integrals in this equation we will solve one by one, starting with the $\vec{p}_T$ integrals.

To be able to make sense of the $\vec{p}_T^{\,2}$ integration in Eq.(II.94) and solve it we have to make two more assumptions in our multiple-splitting model. First, we identify the scale of the strong coupling α_s with the transverse momentum scale of the splitting $\mu_R^2 = \vec{p}_T^{\,2}$. This way we can fully integrate the integrand $\alpha_s/(2\pi)$ and link the final result to the global boundary μ_F.

In addition, we assume strongly ordered splittings in terms of the transverse momentum. If the ordering of the splitting is fixed externally by the chain of momentum fractions x_j, the first splitting, integrated over $\vec{p}_{T,1}^{\,2}$, is now bounded from above by the next external scale $\vec{p}_{T,2}^{\,2}$, which is then bounded by $\vec{p}_{T,3}^{\,2}$, etc. For the n-fold $\vec{p}_T^{\,2}$ integration this means

$$\mu_0^2 < \vec{p}_{T,1}^{\,2} < \vec{p}_{T,2}^{\,2} < \cdots < \mu_F^2 \tag{II.95}$$

At this point this is simply an *ad hoc* assumption which we need to motivate eventually.

Under these three assumptions the transverse momentum integrals in Eq.(II.94) become

$$\begin{aligned}
&\int_{\mu_0}^{\mu_F} \frac{d\vec{p}_{T,n}^2}{\vec{p}_{T,n}^2} \frac{\alpha_s(\vec{p}_{T,n}^2)}{2\pi} \cdots \int_{\mu_0}^{p_{T,3}} \frac{d\vec{p}_{T,2}^2}{\vec{p}_{T,2}^2} \frac{\alpha_s(\vec{p}_{T,2}^2)}{2\pi} \int_{\mu_0}^{p_{T,2}} \frac{d\vec{p}_{T,1}^2}{\vec{p}_{T,1}^2} \frac{\alpha_s(\vec{p}_{T,1}^2)}{2\pi} \cdots \\
&= \int_{\mu_0}^{\mu_F} \frac{d\vec{p}_{T,n}^2}{\vec{p}_{T,n}^2} \frac{1}{2\pi b_0 \log \frac{\vec{p}_{T,n}^2}{\Lambda_{\text{QCD}}^2}} \cdots \int_{\mu_0}^{p_{T,3}} \frac{d\vec{p}_{T,2}^2}{\vec{p}_{T,2}^2} \frac{1}{2\pi b_0 \log \frac{\vec{p}_{T,2}^2}{\Lambda_{\text{QCD}}^2}} \int_{\mu_0}^{p_{T,2}} \frac{d\vec{p}_{T,1}^2}{\vec{p}_{T,1}^2} \frac{1}{2\pi b_0 \log \frac{\vec{p}_{T,1}^2}{\Lambda_{\text{QCD}}^2}} \cdots \\
&= \frac{1}{(2\pi b_0)^n} \int_{\mu_0}^{\mu_F} \frac{d\vec{p}_{T,n}^2}{\vec{p}_{T,n}^2} \frac{1}{\log \frac{\vec{p}_{T,n}^2}{\Lambda_{\text{QCD}}^2}} \cdots \int_{\mu_0}^{p_{T,3}} \frac{d\vec{p}_{T,2}^2}{\vec{p}_{T,2}^2} \frac{1}{\log \frac{\vec{p}_{T,2}^2}{\Lambda_{\text{QCD}}^2}} \int_{\mu_0}^{p_{T,2}} \frac{d\vec{p}_{T,1}^2}{\vec{p}_{T,1}^2} \frac{1}{\log \frac{\vec{p}_{T,1}^2}{\Lambda_{\text{QCD}}^2}} \cdots .
\end{aligned} \tag{II.96}$$

We can solve the individual integrals by switching variables, for example in the last integral

$$\begin{aligned}
\int_{\mu_0}^{p_{T,2}} \frac{d\vec{p}_{T,1}^2}{\vec{p}_{T,1}^2} \frac{1}{\log \frac{\vec{p}_{T,1}^2}{\Lambda_{\text{QCD}}^2}} &= \int_{\log\log \mu_0^2/\Lambda^2}^{\log\log p_{T,2}^2/\Lambda^2} d \log\log \frac{\vec{p}_{T,1}^2}{\Lambda_{\text{QCD}}^2} \qquad \text{with} \qquad \frac{d(ax)}{(ax)\log x} = d\log\log x \\
&= \int_0^{\log\log p_{T,2}^2/\Lambda^2 - \log\log \mu_0^2/\Lambda^2} d\left(\log\log \frac{\vec{p}_{T,1}^2}{\Lambda_{\text{QCD}}^2} - \log\log \frac{\mu_0^2}{\Lambda_{\text{QCD}}^2}\right) \\
&= \log \frac{\log \vec{p}_{T,1}^2/\Lambda_{\text{QCD}}^2}{\log \mu_0^2/\Lambda_{\text{QCD}}^2}\bigg|_0^{\vec{p}_{T,1}^2 \equiv \vec{p}_{T,2}^2} = \log \frac{\log \vec{p}_{T,2}^2/\Lambda_{\text{QCD}}^2}{\log \mu_0^2/\Lambda_{\text{QCD}}^2} .
\end{aligned} \tag{II.97}$$

This gives us for the chain of transverse momentum integrals

$$\begin{aligned}
&\int_0^{p_{T,n} \equiv \mu_F} d\log \frac{\log \vec{p}_{T,n}^2/\Lambda_{\text{QCD}}^2}{\log \mu_0^2/\Lambda_{\text{QCD}}^2} \cdots \int_0^{p_{T,2} \equiv p_{T,3}} d\log \frac{\log \vec{p}_{T,2}^2/\Lambda_{\text{QCD}}^2}{\log \mu_0^2/\Lambda_{\text{QCD}}^2} \int_0^{p_{T,1} \equiv p_{T,2}} d\log \frac{\log \vec{p}_{T,1}^2/\Lambda_{\text{QCD}}^2}{\log \mu_0^2/\Lambda_{\text{QCD}}^2} \\
&= \int_0^{p_{T,n} \equiv \mu_F} d\log \frac{\log \vec{p}_{T,n}^2/\Lambda_{\text{QCD}}^2}{\log \mu_0^2/\Lambda_{\text{QCD}}^2} \cdots \int_0^{p_{T,2} \equiv p_{T,3}} d\log \frac{\log \vec{p}_{T,2}^2/\Lambda_{\text{QCD}}^2}{\log \mu_0^2/\Lambda_{\text{QCD}}^2} \left(\log \frac{\log \vec{p}_{T,2}^2/\Lambda_{\text{QCD}}^2}{\log \mu_0^2/\Lambda_{\text{QCD}}^2}\right) \\
&= \int_0^{p_{T,n} \equiv \mu_F} d\log \frac{\log \vec{p}_{T,n}^2/\Lambda_{\text{QCD}}^2}{\log \mu_0^2/\Lambda_{\text{QCD}}^2} \cdots \frac{1}{2}\left(\log \frac{\log \vec{p}_{T,3}^2/\Lambda_{\text{QCD}}^2}{\log \mu_0^2/\Lambda_{\text{QCD}}^2}\right)^2 \\
&= \int_0^{p_{T,n} \equiv \mu_F} d\log \frac{\log \vec{p}_{T,n}^2/\Lambda_{\text{QCD}}^2}{\log \mu_0^2/\Lambda_{\text{QCD}}^2} \left(\frac{1}{2} \cdots \frac{1}{n-1}\right) \left(\log \frac{\log \vec{p}_{T,n}^2/\Lambda_{\text{QCD}}^2}{\log \mu_0^2/\Lambda_{\text{QCD}}^2}\right)^{n-1} \\
&= \frac{1}{n!}\left(\log \frac{\log \mu_F^2/\Lambda_{\text{QCD}}^2}{\log \mu_0^2/\Lambda_{\text{QCD}}^2}\right)^n = \frac{1}{n!}\left(\log \frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu_F^2)}\right)^n .
\end{aligned} \tag{II.98}$$

This is the final result for the chain of transverse momentum integrals in Eq.(II.94). By assumption, the strong coupling is evaluated at the factorization scale μ_F, which means we identify $\mu_R \equiv \mu_F$.

To compute the convolution integrals over the momentum fractions in Eq.(II.94),

$$\sigma_{n+1}(x,\mu) \sim \frac{1}{n!}\left(\frac{1}{2\pi b_0} \log \frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)}\right)^n \int_{x_0}^1 \frac{dx_n}{x_n} \hat{P}_{g\leftarrow g}\left(\frac{x}{x_n}\right) \cdots \int_{x_0}^1 \frac{dx_1}{x_1} \hat{P}_{g\leftarrow g}\left(\frac{x_2}{x_1}\right) \sigma_1(x_1,\mu_0) , \tag{II.99}$$

we again Mellin transform the equation including the transverse momentum integrals in Eq.(II.98) into moment space

$$\begin{aligned}
\mathcal{M}[\sigma_{n+1}](m,\mu) \sim & \frac{1}{n!}\left(\frac{1}{2\pi b_0}\log\frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)}\right)^n \mathcal{M}\left[\int_{x_0}^1 \frac{dx_n}{x_n}\,\hat{P}_{g\leftarrow g}\left(\frac{x}{x_n}\right)\cdots\int_{x_0}^1\frac{dx_1}{x_1}\,\hat{P}_{g\leftarrow g}\left(\frac{x_2}{x_1}\right)\,\sigma_1(x_1,\mu_0)\right](m) \\
= & \frac{1}{n!}\left(\frac{1}{2\pi b_0}\log\frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)}\right)^n \gamma(m)^n\,\mathcal{M}[\sigma_1](m,\mu_0) \qquad\qquad \text{using } \gamma(m)\equiv\mathcal{M}[P](m) \\
= & \frac{1}{n!}\left(\frac{1}{2\pi b_0}\log\frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)}\,\gamma(m)\right)^n \mathcal{M}[\sigma_1](m,\mu_0)\,. \qquad\qquad \text{(II.100)}
\end{aligned}$$

We can now sum the production cross sections for n collinear jets and obtain

$$\begin{aligned}
\sum_{n=0}^\infty \mathcal{M}[\sigma_{n+1}](m,\mu) = & \mathcal{M}[\sigma_1](m,\mu_0)\,\sum_n \frac{1}{n!}\left(\frac{1}{2\pi b_0}\log\frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)}\,\gamma(m)\right)^n \\
= & \mathcal{M}[\sigma_1](m,\mu_0)\,\exp\left(\frac{\gamma(m)}{2\pi b_0}\,\log\frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)}\right)\,. \qquad\qquad \text{(II.101)}
\end{aligned}$$

This way we can write the Mellin transform of the $(n+1)$ particle production rate as the product of the n-particle rate times a ratio of the strong coupling at two scales

$$\boxed{\sum_{n=0}^\infty \mathcal{M}[\sigma_{n+1}](m,\mu) = \mathcal{M}[\sigma_1](m,\mu_0)\,\left(\frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)}\right)^{\frac{\gamma(m)}{2\pi b_0}}}\,. \qquad \text{(II.102)}$$

This is the same structure as the DGLAP equation's solution in Eq.(II.90). It means that we should be able to understand the physics of the DGLAP equation using our model calculation of a gluon ladder emission.

We should remind ourselves of the three assumptions we need to make to arrive at this form. There are two assumptions which concern the transverse momenta of the successive radiation: first, the global upper limit on all transverse momenta should be the factorization scale μ_F, with a strong ordering in the transverse momenta. This gives us a physical picture of the successive splittings as well as a physical interpretation of the factorization scale. Second, the strong coupling should be evaluated at the transverse momentum or factorization scale, so all scales are unified, in accordance with the derivation of the DGLAP equation.

Bending the rules of pure Yang–Mills QCD we can come back to the hard process σ_1 as the Drell–Yan process $q\bar{q}\to Z$. Each step in n means an additional parton in the final state, so σ_{n+1} is Z production with n collinear partons On the left hand side of Eq.(II.102) we have the sum over any number of additional collinear partons; on the right hand side we see fixed order Drell–Yan production without any additional partons, but with an exponentiated correction factor. Comparing this to the running parton densities we can draw the analogy that any process computed with a scale dependent parton density where the scale dependence is governed by the DGLAP equation includes any number of collinear partons.

We can also identify the logarithms which are resummed by scale dependent parton densities. Going back to Eq.(II.36) reminds us that we start from the divergent collinear logarithms $\log p_T^{\max}/p_T^{\min}$ arising from the collinear phase space integration. In our model for successive splitting we replace the upper boundary by μ_F. The collinear logarithm of successive initial–state parton splitting diverges for $\mu_0\to 0$, but it gets absorbed into the parton densities and determines the structure of the DGLAP equation and its solutions. The upper boundary μ_F tells us to what extent we assume incoming quarks and gluons to be a coupled system of splitting partons and what the maximum momentum scale of these splittings is. Transverse momenta $p_T > \mu_F$ generated by hard parton splitting are not covered by the DGLAP equation and hence not a feature of the incoming partons anymore. They belong to the hard process and have to be consistently simulated. While this scale can be chosen freely we have to make sure that it does not become too large, because at some point the collinear approximation $C\simeq$ constant in Eq.(II.36) ceases to hold and with it our entire argument. Only if we do everything correctly, the DGLAP equation resums logarithms of the maximal transverse momentum size of the incoming gluon. They are universal and arise from simple kinematics.

Looking back at Sections II A and II C we introduce the factorization and renormalization scales step by step completely in parallel: first, computing perturbative higher order contributions to scattering amplitudes we encounter ultraviolet and infrared divergences. We regularize both of them using dimensional regularization with $n=4-2\epsilon<4$ for ultraviolet and $n>4$ for infrared divergences, linked by analytic continuation. Both kinds of divergences are

	renormalization scale μ_R	factorization scale μ_F
source	ultraviolet divergence	collinear (infrared) divergence
poles cancelled	counter terms (renormalization)	parton densities (mass factorization)
summation	resum self energy bubbles	resum parton splittings
parameter	running coupling $\alpha_s(\mu_R^2)$	running parton density $f_j(x,\mu_F)$
evolution	RGE for α_s	DGLAP equation
large scales	decrease of σ_{tot}	increase of σ_{tot} for gluons/sea quarks
theory background	renormalizability	factorization
	proven for gauge theories	proven all orders for DIS
		proven order-by-order DY...

TABLE II: Comparison of renormalization and factorization scales appearing in LHC cross sections.

universal, which means that they are not process or observable dependent. This allows us to absorb ultraviolet and infrared divergences into a re-definition of the strong coupling and the parton density. This nominally infinite shift of parameters we refer to as renormalization for example of the strong coupling or as mass factorization absorbing infrared divergences into the parton distributions.

After renormalization as well as after mass factorization we are left with a scale artifact. Scales arise as part of a the pole subtraction: together with the pole $1/\epsilon$ we have a choice of finite contributions which we subtract with this pole. Logarithms of the renormalization and factorization scales will always be part of these finite terms. Moreover, in both cases the re-definition of parameters is not based on fixed order perturbation theory. Instead, it involves summing logarithms which otherwise can become large and spoil the convergence of our perturbative series in α_s. The only special feature of infrared divergences as compared to ultraviolet divergences is that to identify the resummed logarithms we have to unify both scales to one.

The hadronic production cross section for the Drell–Yan process or other LHC production channels, now including both scales, reads

$$\sigma_{\text{tot}}(\mu_F,\mu_R) = \int_0^1 dx_1 \int_0^1 dx_2 \sum_{ij} f_i(x_1,\mu_F)\, f_j(x_2,\mu_F)\, \hat{\sigma}_{ij}(x_1x_2S,\alpha_s(\mu_R^2),\mu_F,\mu_R)\,. \tag{II.103}$$

The Drell–Yan process has the particular feature that at leading order $\hat{\sigma}_{q\bar{q}}$ only involves weak couplings, it does not include α_s with its implicit renormalization scale dependence at leading order. Strictly speaking, in Eq.(II.103) the parton densities also depend on the renormalization scale because in their extraction we identify both scales. Carefully following their extraction we can separate the two scales if we need to. Lepton pair production and Higgs production in weak boson fusion are the two prominent electroweak production processes at the LHC.

The evolution of all running parameters from one renormalization/factorization scale to another is described either by renormalization group equation in terms of a beta function in the case of renormalization or by the DGLAP equation in the case of mass factorization. Our renormalization group equation for α_s is a single equation, but in general they are sets of coupled differential equations for all relevant parameters, which again makes them more similar to the DGLAP equation.

There is one formal difference between these two otherwise very similar approaches. The fact that we can absorb ultraviolet divergences into process–independent, universal counter terms is called renormalizability and has been proven to all orders for the kind of gauge theories we are dealing with. The universality of infrared splitting kernels has not (yet) in general been proven, but on the other hand we have never seen an example where is fails for sufficiently inclusive observables like production rates. For a while we thought there might be a problem with factorization in supersymmetric theories using the $\overline{\text{MS}}$ scheme, but this issue has been resolved. A summary of the properties of the two relevant scales for LHC physics we show in Table II.

The way we introduce factorization and renormalization scales clearly labels them as an artifact of perturbation theories with divergences. What actually happens if we include all orders in perturbation theory? For example, the resummation of the self energy bubbles simply deals with one class of diagrams which have to be included, either order-by-order or rearranged into a resummation. Once we include all orders in perturbation theory it does not matter according to which combination of couplings and logarithms we order it. An LHC production rate will then not depend on arbitrarily chosen renormalization or factorization scales μ.

Practically, in Eq.(II.103) we evaluate the renormalized parameters and the parton densities at some scale. This scale dependence will only cancel once we include all implicit and explicit appearances of the scales at all orders. Whatever scale we choose for the strong coupling or parton densities will eventually be compensated by explicit scale logarithms. In the ideal case, these logarithms are small and do not spoil perturbation theory. In a process with one distinct external scale, like the Z mass, we know that all scale logarithms should have the form $\log(\mu/m_Z)$. This logarithm vanishes if we evaluate everything at the 'correct' external energy scale, namely m_Z. In that sense we can think of the running coupling as a proper running observable which depends on the external energy of the process. This dependence on the external energy is not a perturbative artifact, because a cross section even to all orders does depend on the energy. The problem in particular for LHC analyses is that after analysis cuts every process will have more than one external energy scale.

We can turn around the argument of vanishing scale dependence to all orders in perturbation theory. This gives us an estimate of the minimum theoretical error on a rate prediction set by the scale dependence. The appropriate interval of what we consider reasonable scale choices depends on the process and the taste of the people doing this analysis. This error estimate is not at all conservative; for example the renormalization scale dependence of the Drell–Yan production rate or Higgs production in weak boson fusion is zero because α_s only enters are next–to–leading order. At the same time we know that the next–to–leading order correction to the Drell–Yan cross section is of the order of 30%, which far exceeds the factorization scale dependence. Moreover, the different scaling behavior of a hadronic cross section shown in Table II implies that for example gluon–induced processes at typical x values around 10^{-2} show a cancellation of the factorization and renormalization scale variation. Estimating theoretical uncertainties from scale dependence therefore requires a good understanding of the individual process and the way it is affected by the two scales.

Guessing the right scale choice for a process is hard, often impossible. For example in Drell–Yan production at leading order there exists only one scale, m_Z. If we set $\mu = m_Z$ all scale logarithms vanish. In reality, LHC observables include several different scales. Some of them appear in the hard process, for example in the production of two or three particles with different masses. Others enter through the QCD environment where at the LHC we only consider final–state jets above a certain minimal transverse momentum. Even others appear though background rejection cuts in a specific analysis, for example when we only consider the Drell–Yan background for $m_{\mu\mu} > 1$ TeV to Kaluza–Klein graviton production. Using likelihood methods does not improve the situation because the phase space regions dominated by the signal will introduce specific energy scales which affect the perturbative prediction of the backgrounds. This is one of the reasons why an automatic comparison of LHC events with signal or background predictions is bound to fail once it requires an estimate of the theoretical uncertainty on the background simulation.

All that means that in practice there is no way to define a 'correct' scale. On the other hand, there are definitely poor scale choices. For example, using $1000 \times m_Z$ as a typical scale in the Drell–Yan process will if nothing else lead to logarithms of the size $\log 1000$ whenever a scale logarithm appears. These logarithms eventually have to be cancelled to all orders in perturbation theory, inducing unreasonably large higher order corrections.

When describing jet radiation, we usually introduce a phase space dependent renormalization scale, evaluating the strong coupling at the transverse momentum of the radiated jet $\alpha_s(\vec{p}^{\,2}_{T,j})$. This choice gives the best kinematic distributions for the additional partons because in Section II E we have shown that it resums large collinear logarithms.

The transverse momentum of a final–state particle is one of scale choices allowed by factorization; in addition to poor scale choices there also exist wrong scale choices, *i.e.* scale choices violating physical properties we need. Factorization or the Kinoshita–Lee–Nauenberg theorem which ensures that soft divergences cancel between real and virtual emission diagrams are such properties we should not violate — in QED the same property is called the Bloch–Nordsieck cancellation. Imagine picking a factorization scale defined by the partonic initial state, for example the partonic center–of–mass energy $s = x_1 x_2 S$. We know that this definition is not unique: for any final state it corresponds to the well defined sum of all momenta squared. However, virtual and real gluon emission generate different multiplicities in the final state, which means that the two sources of soft divergences only cancel until we multiply each of them with numerically different parton densities. Only scales which are uniquely defined in the final state can serve as factorization scales. For the Drell–Yan process such a scale could be m_Z, or the mass of heavy new-physics states in their production process. So while there is no such thing as a correct scale choice, there are more or less smart choices, and there are definitely very poor choices, which usually lead to an unstable perturbative behavior.

F. Parton shower

In LHC phenomenology we are usually less interested in fixed-order perturbation theory than in logarithmically enhanced QCD effects, for example collinear jet radiation. After introducing the kernels $P_{i \leftarrow j}(z)$ as something like

splitting probabilities we never applied a probabilistic approach to parton splitting. The basis of such an interpretation are Sudakov factors describing the splitting of a parton i into any of the partons j based on the factorized form Eq.(II.48)

$$\boxed{\Delta_i(t) \equiv \Delta_i(t,t_0) = \exp\left(-\sum_j \int_{t_0}^{t} \frac{dt'}{t'} \int_0^1 dy\ \frac{\alpha_s}{2\pi} \hat{P}_{j\leftarrow i}(y)\right)} \,. \tag{II.104}$$

We can express all Sudakov factors in terms of splitting functions Γ_j,

$$\begin{aligned}
\Delta_q(t) &= \exp\left(-\int_{t_0}^{t} dt'\ \Gamma_{q\leftarrow q}(t,t')\right) \\
\Delta_g(t) &= \exp\left(-\int_{t_0}^{t} dt'\ \left[\Gamma_{g\leftarrow g}(t,t') + \Gamma_{q\leftarrow g}(t')\right]\right) \,,
\end{aligned} \tag{II.105}$$

which to leading logarithm in $\log(t/t')$ read

$$\begin{aligned}
\Gamma_{q\leftarrow q}(t,t') &= \frac{C_F}{2\pi}\ \frac{\alpha_s(t')}{t'}\left(\log\frac{t}{t'} - \frac{3}{2}\right) \\
\Gamma_{g\leftarrow g}(t,t') &= \frac{C_A}{2\pi}\ \frac{\alpha_s(t')}{t'}\left(\log\frac{t}{t'} - \frac{11}{6}\right) \\
\Gamma_{q\leftarrow g}(t') &= \frac{n_f}{6\pi}\ \frac{\alpha_s(t')}{t'} \,.
\end{aligned} \tag{II.106}$$

These formulas have a slight problem: terms arising from next–to–leading logarithms spoil the limit $t' \to t$, where a splitting probability should vanish. Technically, we can deal with the finite terms in the Sudakov factors by requiring them to be positive semi–definite, *i.e.* by replacing $\Gamma(t,t') < 0$ with zero. For the general argument this problem with the analytic expressions for the splitting functions is irrelevant. In practice, many modern parton showers do not use these approximate formulas and instead integrate the full splitting kernels.

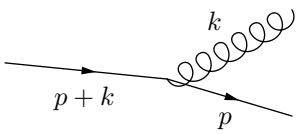
Instead of going in the details of Sudakov factors we only sketch how such exponentials appear in probabilistic arguments. Using Poisson statistics for something expected to occur $\langle n\rangle$ times, the probability of observing it n times is given by

$$\mathcal{P}(n;\langle n\rangle) = \frac{\langle n\rangle^n\, e^{-\langle n\rangle}}{n!} \qquad\qquad \mathcal{P}(0;\langle n\rangle) = e^{-\langle n\rangle} \,. \tag{II.107}$$

If the exponent in the Sudakov factor describes the integrated splitting probability of a parton i this means that the Sudakov itself describes a non–splitting probability of the parton i into any final state j. In other words, we can use splitting probabilities to describe initial state radiation and final state radiation, including a well-defined probability of no radiation at all.

While the Sudakov factors and splitting probabilities allow us to describe for example the emission of gluons off a hard quark line, there is a crucial piece missing from our collinear resummation argument: multiple gluon emission has to be ordered, and there cannot be interference terms between different emission stages. Such interference diagrams contributing to the full amplitude squared are called non–planar diagrams. There are three reasons to neglect them, even though none of them gives exactly zero for soft and collinear splittings. On the other hand, in combination they make for a very good reason.

First, an arguments for a strongly ordered gluon emission comes from the divergence structure of soft and collinear gluon emission. We start with the kinematical setup for gluon radiation off a massless or massive hard quark leg



The original massive quark leg with momentum $p + k$ and mass m could be attached to some hard process as a splitting final state. It splits into a hard quark p and a soft gluon k. The general matrix element without any

approximation reads

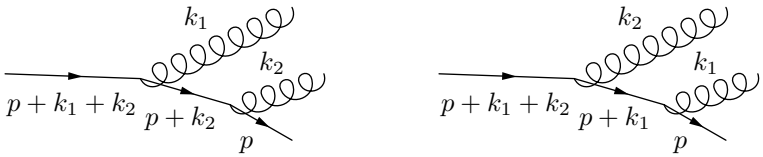
$$\begin{aligned}\mathcal{M}_{n+1} &= g_s T^a\, \epsilon_\mu^*(k)\, \bar{u}(p)\gamma^\mu \frac{\not{p}+\not{k}+m}{(p+k)^2-m^2}\, \mathcal{M}_n \\ &= g_s T^a\, \epsilon_\mu^*(k)\, \bar{u}(p)\left[-\not{p}\gamma^\mu + 2p^\mu + m\gamma^\mu + \gamma^\mu \not{k}\right]\, \frac{1}{2(pk)+k^2}\, \mathcal{M}_n \\ &= g_s T^a\, \epsilon_\mu^*(k)\, \bar{u}(p) \frac{2p^\mu + \gamma^\mu \not{k}}{2(pk)+k^2}\, \mathcal{M}_n\,, \end{aligned} \tag{II.108}$$

using the Dirac equation $\bar{u}(p)(\not{p}-m)=0$. At this level, a further simplification requires for example the soft gluon limit. In the presence of only hard momenta, except for the gluon, we can define it for example as $k^\mu = \lambda p^\mu$, where p^μ is an arbitrary four-vector combination of the surrounding hard momenta. The small parameter λ then characterizes the soft limit. For the invariant mass of the gluon we assume $k^2 = \mathcal{O}(\lambda^2)$, allowing for a slightly off–shell gluon. We find the leading contribution

$$\mathcal{M}_{n+1} \sim g_s T^a\, \epsilon_\mu^*(k)\, \frac{p^\mu}{(pk)}\, \bar{u}(p)\, \mathcal{M}_n\,. \tag{II.109}$$

To allow for other color states we defines a color operator $\hat{T}_j^a$ which we insert into the matrix element of Eq.(II.109) and which assumes values of $+T_{ij}^a$ for radiation off a quark, $-T_{ji}^a$ for radiation off an antiquark and $-if_{abc}$ for radiation off a gluon. For a color neutral process like our favorite Drell–Yan process adding an additional soft gluon $q\bar{q} \to Zg$ we find for the sum of all possible soft emissions $\sum_j \hat{T}_j = 0$.

Radiating two radiated gluons instead of one look like



Each soft splitting comes with a term $(\epsilon^* p)(pk)$, so the two Feynman diagrams give us the combined kinetic terms

$$\begin{aligned}&\frac{(\epsilon_1 p)}{(p+k_1+k_2)^2-m^2}\frac{(\epsilon_2 p)}{(p+k_2)^2-m^2} + \frac{(\epsilon_2 p)}{(p+k_1+k_2)^2-m^2}\frac{(\epsilon_1 p)}{(p+k_1)^2-m^2} \\ &\quad= \frac{(\epsilon_1 p)}{2(pk_1)+2(pk_2)+(k_1+k_2)^2}\frac{(\epsilon_2 p)}{2(pk_2)} + \frac{(\epsilon_2 p)}{2(pk_1)+2(pk_2)+(k_1+k_2)^2}\frac{(\epsilon_1 p)}{2(pk_1)} \qquad k_1^2=0=k_2^2 \\ &\quad\simeq \frac{(\epsilon_1 p)}{2\max_j(pk_j)}\frac{(\epsilon_2 p)}{2(pk_2)} + \frac{(\epsilon_2 p)}{2\max_j(pk_j)}\frac{(\epsilon_1 p)}{2(pk_1)} \qquad (pk_j) \text{ strongly ordered} \\ &\quad\simeq \begin{cases} \dfrac{(\epsilon_1 p)(\epsilon_2 p)}{2\max_j(pk_j)}\,\dfrac{1}{2(pk_2)} & (pk_2) \ll (pk_1)\;\; k_2 \text{ softer} \\ \dfrac{(\epsilon_1 p)(\epsilon_2 p)}{2\max_j(pk_j)}\,\dfrac{1}{2(pk_1)} & (pk_1) \ll (pk_2)\;\; k_1 \text{ softer}\,. \end{cases}\end{aligned} \tag{II.110}$$

For the two Feynman diagrams this means that once one of the gluons is significantly softer than the other the Feynman diagrams with the later soft emission dominates. This argument can be generalized to multiple gluon emission by recognizing that the kinematics will always be dominated by the more divergent propagators towards the final state quark with momentum p. Note, however, that it is based on an ordering of the scalar products (pk_j) interpreted as the softness of the gluons, and we already know that a small value of (pk_j) can as well point to a collinear divergence.

Second, we can derive ordered multiple gluon emission from the phase space integration in the soft limit. For this we need to square this matrix element of Eq.(II.109). It includes a polarization sum and will therefore depend on the gauge. We choose the general axial gauge for massless gauge bosons

$$\sum_{\text{pol}} \epsilon_\mu^*(k)\epsilon_\nu(k) = -g_{\mu\nu} + \frac{k_\mu n_\nu + n_\mu k_\nu}{(nk)} - n^2\frac{k_\mu k_\nu}{(nk)^2} = -g_{\mu\nu} + \frac{k_\mu n_\nu + n_\mu k_\nu}{(nk)}\,, \tag{II.111}$$

with a light-like reference vector n obeying $n^2 = 0$. The matrix element squared then reads

$$\begin{aligned}
\overline{|\mathcal{M}_{n+1}|^2} &= g_s^2 \left(-g_{\mu\nu} + \frac{k_\mu n_\nu + n_\mu k_\nu}{(nk)}\right) \left(\sum_j \hat{T}_j^a \frac{p_j^\mu}{(p_j k)}\right)^\dagger \left(\sum_j \hat{T}_j^a \frac{p_j^\nu}{(p_j k)}\right) \overline{|\mathcal{M}_n|^2} \\
&= g_s^2 \left(-\left(\sum_j \hat{T}_j^a \frac{p_j^\mu}{(p_j k)}\right)^\dagger \left(\sum_j \hat{T}_j^a \frac{p_{j\mu}}{(p_j k)}\right) + \frac{2}{(nk)} \left(\sum_j \hat{T}_j^a\right)^\dagger \left(\sum_j \hat{T}_j^a \frac{(p_j n)}{(p_j k)}\right)\right) \overline{|\mathcal{M}_n|^2} \\
&= -g_s^2 \left(\sum_j \hat{T}_j^a \frac{p_j^\mu}{(p_j k)}\right)^\dagger \left(\sum_j \hat{T}_j^a \frac{p_{j\mu}}{(p_j k)}\right) \overline{|\mathcal{M}_n|^2} \,. \qquad \text{(II.112)}
\end{aligned}$$

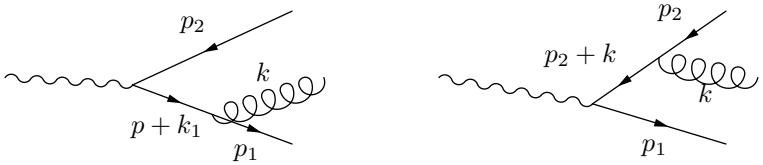
The insertion operator in the matrix element has the form of an insertion current multiplied by its hermitian conjugate. This current describes the universal form of soft gluon radiation off an n-particle process

$$\boxed{\overline{|\mathcal{M}_{n+1}|^2} \equiv -g_s^2 \, (J^\dagger \cdot J) \, \overline{|\mathcal{M}_n|^2}} \qquad \text{with} \quad J^{a\mu}(k, \{p_j\} = \sum_j \hat{T}_j^a \, \frac{p_j^\mu}{(p_j k)} \,. \qquad \text{(II.113)}$$

We can further simplify the squared current to

$$\begin{aligned}
(J^\dagger \cdot J) &= \sum_j \hat{T}_j^a \hat{T}_j^a \, \frac{p_j^2}{(p_j k)^2} + 2 \sum_{i<j} \hat{T}_i^a \hat{T}_j^a \, \frac{(p_i p_j)}{(p_i k)(p_j k)} \\
&= \sum_j \hat{T}_j^a \left(-\sum_{i \neq j} \hat{T}_i^a\right) \frac{p_j^2}{(p_j k)^2} + 2 \sum_{i<j} \hat{T}_i^a \hat{T}_j^a \, \frac{(p_i p_j)}{(p_i k)(p_j k)} \\
&= -\left(\sum_{i<j} + \sum_{i>j}\right) \hat{T}_i^a \hat{T}_j^a \frac{p_j^2}{(p_j k)^2} + 2 \sum_{i<j} \hat{T}_i^a \hat{T}_j^a \, \frac{(p_i p_j)}{(p_i k)(p_j k)} \\
&= 2 \sum_{i<j} \hat{T}_i^a \hat{T}_j^a \left(\frac{(p_i p_j)}{(p_i k)(p_j k)} - \frac{p_i^2}{2(p_i k)^2} - \frac{p_j^2}{2(p_j k)^2}\right) \qquad \text{in the general massive case} \\
&= 2 \sum_{i<j} \hat{T}_i^a \hat{T}_j^a \, \frac{(p_i p_j)}{(p_i k)(p_j k)} \qquad \text{for massless partons .} \qquad \text{(II.114)}
\end{aligned}$$

A simple process which will illustrate the relevant point for multiple gluon emission is



We first symmetrize the leading soft radiation term with respect to the two hard momenta in a particular way,

$$\begin{aligned}
(J^\dagger \cdot J)_{12} &= \frac{(p_1 p_2)}{(p_1 k)(p_2 k)} \\
&= \frac{1}{k_0^2} \frac{1 - \cos\theta_{12}}{(1 - \cos\theta_{1k})(1 - \cos\theta_{2k})} \qquad \text{in terms of opening angles } \theta \\
&= \frac{1}{2k_0^2} \left(\frac{1 - \cos\theta_{12}}{(1 - \cos\theta_{1k})(1 - \cos\theta_{2k})} + \frac{1}{1 - \cos\theta_{1k}} - \frac{1}{1 - \cos\theta_{2k}}\right) + (1 \leftrightarrow 2) \\
&\equiv \frac{W_{12}^{[1]} + W_{12}^{[2]}}{k_0^2} \,. \qquad \text{(II.115)}
\end{aligned}$$

The last term is an implicit definition of the two terms $W_{12}^{[1]}$. The pre-factor $1/k_0^2$ is given by the leading soft

divergence. The original form of $(J^\dagger J)$ is symmetric in the two indices, which means that both hard partons can take the role of the hard parton and the interference partner. In the new form the symmetry in each of the two terms is broken. Each of the two terms we need to integrate over the gluon's phase space, including the azimuthal angle ϕ_{1k}.

To compute the actual integral we express the three parton vectors in polar coordinates where the initial parton p_1 propagates into the x direction, the interference partner p_2 in the $(x-y)$ plane, and the soft gluon in the full three-dimensional space described by polar coordinates,

$$\begin{aligned} \hat{p}_1 &= (1,0,0) && \text{hard parton} \\ \hat{p}_2 &= (\cos\theta_{12}, \sin\theta_{12}, 0) && \text{interference partner} \\ \hat{k} &= (\cos\theta_{1k}, \sin\theta_{1k}\cos\phi_{1k}, \sin\theta_{1k}\sin\phi_{1k}) && \text{soft gluon} \\ \Rightarrow \qquad \cos\theta_{2k} &\equiv (\hat{p}_2\hat{k}) = \cos\theta_{12}\cos\theta_{1k} + \sin\theta_{12}\sin\theta_{1k}\cos\phi_{1k}\,. && \end{aligned} \tag{II.116}$$

From the scalar product between these four-vectors we see that of the terms appearing in Eq.(II.115) only the opening angle θ_{2k} includes ϕ_{1k}, which for the azimuthal angle integration means

$$\begin{aligned} \int_0^{2\pi} d\phi_{1k}\, W_{12}^{[1]} &= \frac{1}{2}\int_0^{2\pi} d\phi_{1k}\, \left(\frac{1-\cos\theta_{12}}{(1-\cos\theta_{1k})(1-\cos\theta_{2k})} + \frac{1}{1-\cos\theta_{1k}} - \frac{1}{1-\cos\theta_{2k}} \right) . \\ &= \frac{1}{2}\,\frac{1}{1-\cos\theta_{1k}} \int_0^{2\pi} d\phi_{1k}\, \left(\frac{1-\cos\theta_{12}}{1-\cos\theta_{2k}} + 1 - \frac{1-\cos\theta_{1k}}{1-\cos\theta_{2k}} \right) \\ &= \frac{1}{2}\,\frac{1}{1-\cos\theta_{1k}} \left(2\pi + (\cos\theta_{1k} - \cos\theta_{12}) \int_0^{2\pi} d\phi_{1k}\, \frac{1}{1-\cos\theta_{2k}} \right) . \end{aligned} \tag{II.117}$$

The azimuthal angle integral in this expression for $W_{12}^{[i]}$ we can solve

$$\begin{aligned} \int_0^{2\pi} d\phi_{1k} \frac{1}{1-\cos\theta_{2k}} &= \int_0^{2\pi} d\phi_{1k} \frac{1}{1-\cos\theta_{12}\cos\theta_{1k} - \sin\theta_{12}\sin\theta_{1k}\cos\phi_{1k}} \\ &= \int_0^{2\pi} d\phi_{1k} \frac{1}{a - b\cos\phi_{1k}} \\ &= \oint_{\text{unit circle}} dz\, \frac{1}{iz} \frac{1}{a - b\dfrac{z+1/z}{2}} && \text{with} \quad z = e^{i\phi_{1k}},\ \cos\phi_{1k} = \frac{z+1/z}{2} \\ &= \frac{2}{i} \oint dz\, \frac{1}{2az - b - bz^2} \\ &= \frac{2i}{b} \oint \frac{dz}{(z-z_-)(z-z_+)} && \text{with} \quad z_\pm = \frac{a}{b} \pm \sqrt{\frac{a^2}{b^2} - 1}\,. \end{aligned} \tag{II.118}$$

This integral is related to the sum of all residues of poles inside the closed integration contour. Of the two poles z_- is the one which typically lies within the unit circle, so we find

$$\begin{aligned} \int_0^{2\pi} d\phi_{1k} \frac{1}{1-\cos\theta_{2k}} &= \frac{2i}{b}\, 2\pi i \frac{1}{z_- - z_+} = \frac{2\pi}{\sqrt{a^2-b^2}} \\ &= \frac{2\pi}{\sqrt{(\cos\theta_{1k} - \cos\theta_{12})^2}} = \frac{2\pi}{|\cos\theta_{1k} - \cos\theta_{12}|}\,. \end{aligned} \tag{II.119}$$

The entire integral in Eq.(II.117) then becomes

$$\begin{aligned} \int_0^{2\pi} d\phi_{1k}\, W_{12}^{[1]} &= \frac{1}{2}\,\frac{1}{1-\cos\theta_{1k}} \left(2\pi + (\cos\theta_{1k} - \cos\theta_{12})\, \frac{2\pi}{|\cos\theta_{1k} - \cos\theta_{12}|} \right) \\ &= \frac{\pi}{1-\cos\theta_{1k}} \left(1 + \text{sign}(\cos\theta_{1k} - \cos\theta_{12})\right) \\ &= \begin{cases} \dfrac{2\pi}{1-\cos\theta_{1k}} & \text{if} \quad \theta_{1k} < \theta_{12} \\ 0 & \text{else}\,. \end{cases} \end{aligned} \tag{II.120}$$

The soft gluon is only radiated at angles between zero and the opening angle of the initial parton p_1 and its hard interference partner or spectator p_2. The same integral over $W_{12}^{[2]}$ gives the same result, with switched roles of p_1 and p_2. Combining the two permutations this means that the soft gluon is always radiated within a cone centered around one of the hard partons and with a radius given by the distance between the two hard partons. The coherent sum of diagrams reduces to an incoherent sum. This derivation angular ordering is exact in the soft limit.

There is a simple physical argument for this suppressed radiation outside a cone defined by the radiating legs. Part of the deviation is that the over–all process is color–neutral. This means that once the gluon is far enough from the two quark legs it will not resolve their individual charges but only feel the combined charge. This screening leads to an additional suppression factor of the kind $\theta_{12}^2/\theta_{1k}^2$. This effect is called coherence.

The third argument for ordered emission comes from color factors. Crossed successive splittings or interference terms between different orderings are color suppressed. For example in the squared diagram for three jet production in e^+e^- collisions the additional gluon contributes a color factor

$$\mathrm{tr}(T^a T^a) = \frac{N_c^2 - 1}{2} = N_c C_F \tag{II.121}$$

When we consider the successive radiation of two gluons the ordering matters. As long as the gluon legs do not cross each other we find the color factor

$$\begin{aligned}
\mathrm{tr}(T^a T^a T^b T^b) &= (T^a T^a)_{il} (T^b T^b)_{li} \\
&= \frac{1}{4} \left(\delta_{il}\delta_{jj} - \frac{\delta_{ij}\delta_{jl}}{N_c} \right) \left(\delta_{il}\delta_{jj} - \frac{\delta_{ij}\delta_{jl}}{N_c} \right) \qquad \text{using} \quad T^a_{ij} T^a_{kl} = \frac{1}{2} \left(\delta_{il}\delta_{jk} - \frac{\delta_{ij}\delta_{kl}}{N_c} \right) \\
&= \frac{1}{4} \left(\delta_{il} N_c - \frac{\delta_{il}}{N_c} \right) \left(\delta_{il} N_c - \frac{\delta_{il}}{N_c} \right) \\
&= N_c \left(\frac{N_c^2 - 1}{2N_c} \right)^2 = N_c C_F^2 = \frac{16}{3}
\end{aligned} \tag{II.122}$$

Similarly, we can compute the color factor when the two gluon lines cross. We find

$$\mathrm{tr}(T^a T^b T^a T^b) = -\frac{N_c^2 - 1}{4N_c} = -\frac{C_F}{2} = -\frac{2}{3} \,. \tag{II.123}$$

Numerically, this color factor is suppressed compared to 16/3. This kind of behavior is usually quoted in powers of N_c where we assume N_c to be large. In those terms non–planar diagrams are suppressed by a factor $1/N_c^2$ compared to the planar diagrams.

Once we also include the triple gluon vertex we can radiate two gluons off a quark leg with the color factor

$$\mathrm{tr}(T^a T^b)\; f^{acd} f^{bcd} = \frac{\delta^{ab}}{2}\; N_c \delta^{ab} = \frac{N_c (N_c^2 - 1)}{2} = N_c^2 C_F = \frac{36}{3} \,. \tag{II.124}$$

This is not suppressed compared to successive planar gluon emission, neither in actual numbers not in the large-N_c limit.

We can try the same argument for a purely gluonic theory, *i.e.* radiating gluons off two hard gluons in the final state. The color factor for single gluon emission after squaring is

$$f^{abc} f^{abc} = N_c \delta^{aa} = N_c (N_c^2 - 1) \sim N_c^3 \,, \tag{II.125}$$

using the large-N_c limit in the last step. For planar double gluon emission with the exchanged gluon indices b and f we find

$$f^{abd} f^{abe} f^{dfg} f^{efg} = N_c \delta^{de}\; N_c \delta^{de} \sim N_c^4 \,. \tag{II.126}$$

Splitting one radiated gluon into two gives

$$f^{abc}\; f^{cef} f^{def}\; f^{abd} = N_c \delta^{cd}\; N_c \delta^{cd} \sim N_c^4 \,. \tag{II.127}$$

This means that planar emission of two gluons and successive splitting of one into two gluons cannot be separated

based on the color factor. We can use the color factor argument only for abelian splittings to justify ordered gluon emission.

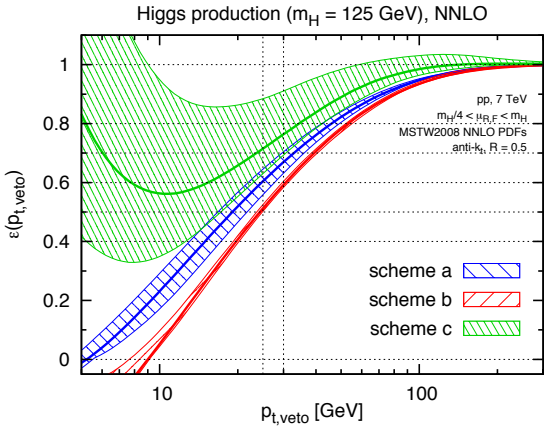


FIG. 4: Different predictions for the jet veto survival probability P_{pass} as a function of the maximum allowed $p_{T,j}$. The example process chosen is Higgs production in gluon fusion. The shaded regions indicate the independent variation of the factorization and renormalization scales within $[m_H/4, m_H]$ requiring μ_R/μ_F to lie within $[0.5, 2]$. The figure and the corresponding physics argument are taken from Ref. [**?**].

III. JETS

Until now we have described QCD in terms of its fundamental constituents, the quarks and the gluons. This framework is the same for all applications of QCD, including collider physics. On the other hand, from the running of the strong coupling we already know that a perturbative theory of quarks and gluons will at some point hit its self-consistency conditions. At colliders this defines the transition from partons to jets, where jets are defined as a spray of particles coming from a highly energetic produced parton. In this chapter we will start from quarks and gluons and slowly transition to jets at the LHC and some very recent developments in this direction.

A. Jet counting

In Higgs physics and other QCD applications the number of jets radiated from the initial (and final) state is an important observable which separates signals from backgrounds. From our discussion of collinear divergences we know that any perturbative calculation for a hadron collider by definition includes any number of collinear jets from initial state splittings. This means that the number of jets in an event in perturbatively very hard to compute: on the one hand, for a given hard process the number of jets is defined through the process; on the other hand, any number of collinear jets is always included, but only in the collinear and hence unobservable phase space configuration.

Experimentally, we can assign a probability pattern to the radiation of jets from the core process. As an example we assume NNLO or two-loop precision for an LHC production rate

$$\sigma = \sigma_0 + \alpha_s \sigma_1 + \alpha_s^2 \sigma_2 \; , \tag{III.1}$$

where we omit the over–all power of α_s in σ_0. Consequently, we define the cross section passing a cut on the number of jets or a jet veto as

$$\sigma^{(\text{pass})} \equiv P_{\text{pass}} \; \sigma = \sum_j \alpha_s^j \sigma_j^{(\text{pass})} \; . \tag{III.2}$$

This relation defines a survival probability P_{pass} of a jet veto. If we assume that the leading order prediction only includes a Higgs or a lepton pair in the final state, we know that $\sigma_0^{(\text{pass})} = \sigma_0$. Perturbatively we can compute the jet veto survival probability as

$$P_{\text{pass}}^{(\text{a})} = \frac{\sigma^{(\text{pass})}}{\sigma} = \frac{\sigma_0 + \alpha_s \sigma_1^{(\text{pass})} + \alpha_s^2 \sigma_2^{(\text{pass})}}{\sigma_0 + \alpha_s \sigma_1 + \alpha_s^2 \sigma_2} \; , \tag{III.3}$$

The result as a function of the maximum allowed $p_{T,j}$ is shown as 'scheme a' in Figure 4. The shaded region is an estimate of the theoretical uncertainty of this prediction.

Alternatively, we can argue that the proper perturbative observable is the fraction of vetoed events $(1 - P_{\text{pass}})$. Indeed, for small values of α_s the jet radiation probability vanishes and with it $(1 - P_{\text{pass}}) \sim \alpha_s \to 0$. This vetoed event fraction we can compute as $\sigma_j - \sigma_j^{(\text{pass})}$ for $j \geq 0$. However, we need to keep in mind that in the presence of an additional hard jet the NNLO rate prediction reduces to NLO accuracy of the inclusive process, so we include only the two leading terms in the numerator and denominator,

$$\begin{aligned} 1 - P_{\text{pass}}^{(\text{b})} &= \alpha_s \, \frac{\sigma_1 - \sigma_1^{(\text{pass})} + \alpha_s(\sigma_2 - \sigma_2^{(\text{pass})})}{\sigma_0 + \alpha_s \sigma_1} \\ P_{\text{pass}}^{(\text{b})} &= 1 - \alpha_s \, \frac{\sigma_1 - \sigma_1^{(\text{pass})} + \alpha_s(\sigma_2 - \sigma_2^{(\text{pass})})}{\sigma_0 + \alpha_s \sigma_1} = \frac{\sigma_0 + \alpha_s \sigma_1^{(\text{pass})} + \alpha_s^2 \sigma_2^{(\text{pass})} - \alpha_s^2 \sigma_2}{\sigma_0 + \alpha_s \sigma_1} \, . \end{aligned} \tag{III.4}$$

This defines 'scheme b' in Figure 4. Obviously, in Eq.(III.4) we can move the term $-\alpha_s^2 \sigma_2$ into the denominator and arrive at Eq.(III.3) within the uncertainty defined by the unknown α_s^3 terms.

Finally, we can consistently expand the definition of P_{pass} as the ratio given in Eq.(III.3). The two leading derivatives of a ratio read

$$\begin{aligned} \left(\frac{f}{g}\right)' &= \frac{f'g - fg'}{g^2} \overset{f=g}{=} \frac{f' - g'}{g} \\ \left(\frac{f}{g}\right)'' &= \left(\frac{f'g}{g^2} - \frac{fg'}{g^2}\right)' = \frac{(f'g)'g^2 - f'g2gg'}{g^4} - \frac{(fg')'g^2 - fg'2gg'}{g^4} \\ &= \frac{(f'g)' - 2f'g'}{g^2} - \frac{(fg')'g - 2fg'g'}{g^3} = \frac{f''g - f'g'}{g^2} - \frac{fg''g - fg'g'}{g^3} \overset{f=g}{=} \frac{f'' - g''}{g} - \frac{g'(f' - g')}{g^2} \end{aligned} \tag{III.5}$$

In the last steps we assume $f = g$ at the point where we evaluate the Taylor expansion. Applied to the perturbative QCD series for $(1 - P_{\text{pass}})$ around the zero-coupling limit this gives us

$$\begin{aligned} 1 - P_{\text{pass}}^{(\text{c})} &= 1 - \frac{\sigma_0 + \alpha_s \sigma_1^{(\text{pass})} + \alpha_s^2 \sigma_2^{(\text{pass})} + \cdots}{\sigma_0 + \alpha_s \sigma_1 + \alpha_s^2 \sigma_2 + \cdots} \\ P_{\text{pass}}^{(\text{c})} &= 1 + \alpha_s \, \frac{\sigma_1^{(\text{pass})} - \sigma_1}{\sigma_0} + \alpha_s^2 \, \frac{\sigma_2^{(\text{pass})} - \sigma_2}{\sigma_0} - \alpha_s^2 \, \frac{\sigma_1(\sigma_1^{(\text{pass})} - \sigma_1)}{\sigma_0^2} \, , \end{aligned} \tag{III.6}$$

defining 'scheme c' in Figure 4. The numerical results indicate that the three schemes are inconsistent within their theoretical uncertainties, and that the most consistent Taylor expansion around perfect veto survival probabilities is doing particularly poorly. Towards small $p_{T,j}$ veto ranges the fixed order perturbative approach clearly fails.

A first, non-perturbative ansatz for the calculation of the number of radiated jets is motivated by soft photon or gluon emission. In the eikonal approximation we know that the number of successively radiated jets is a Poisson distribution in the numbers of jets. If we assume such a Poisson distribution, the probability of observing exactly n jets given an expected $\langle n \rangle$ jets is

$$\mathcal{P}(n; \langle n \rangle) = \frac{\langle n \rangle^n e^{-\langle n \rangle}}{n!} \qquad \Rightarrow \qquad \boxed{P_{\text{pass}} \equiv \mathcal{P}(0; \langle n \rangle) = e^{-\langle n \rangle}} \, . \tag{III.7}$$

This probability links rates for exactly n jets, no at least n jets, *i.e.* it described the exclusive number of jets. A Poisson distribution is normalized to unity, once we sum over all possible jet multiplicities n. We consistently fix the expectation value in terms of the inclusive cross sections producing at least zero or at least one jet,

$$\langle n \rangle = \frac{\sigma_1(p_T^{\text{min}})}{\sigma_0} \, . \tag{III.8}$$

We include an infra-red cut-off p_T^{min}, defining a cross section in the presence of soft and collinear divergences. The inclusive jet ratio σ_1/σ_0 is reproduced by the ratio of the corresponding Poisson distributions. Including this expectation value $\langle n \rangle$ into Eq.(III.7) returns a veto survival probability of $\exp(-\sigma_1/\sigma_0)$.

A second ansatz comes from UA1 and UA2 data, 'phenomenologically' assuming a constant probability of radiating

a jet. In terms of the inclusive cross sections σ_n, *i.e.* the production rate for the radiation of at least n jets, this implies

$$\frac{\sigma_{n+1}(p_T^{\min})}{\sigma_n(p_T^{\min})} = R^{(\text{incl})}_{(n+1)/n}(p_T^{\min}) \; . \tag{III.9}$$

The expected number of jets is then given by

$$\begin{aligned}\langle n\rangle &= \frac{1}{\sigma_0}\sum_{j=1} j(\sigma_j - \sigma_{j+1}) = \frac{1}{\sigma_0}\left(\sum_{j=1} j\sigma_j - \sum_{j=2}(j-1)\sigma_j\right) = \frac{1}{\sigma_0}\sum_{j=1}\sigma_j \\ &= \frac{\sigma_1}{\sigma_0}\sum_{j=0}(R^{(\text{incl})}_{(n+1)/n})^j = \frac{R^{(\text{incl})}_{(n+1)/n}}{1-R^{(\text{incl})}_{(n+1)/n}} \; ,\end{aligned} \tag{III.10}$$

if $R^{(\text{incl})}_{(n+1)/n}$ is a constant. Radiating jets with such a constant probability has been observed at many experiments, including most recently the LHC, and is in the context of W+jets referred to as staircase scaling. The defining property of staircase scaling is that ratios of the $(n+1)$-jet rate to the n-jet rate for inclusive and exclusive jet rates are identical,

$$\begin{aligned}R^{(\text{incl})}_{(n+1)/n} &= \frac{\sigma_{n+1}}{\sigma_n} = \frac{\sum_{j=n+1}^{\infty}\sigma_j^{(\text{excl})}}{\sigma_n^{(\text{excl})} + \sum_{j=n+1}^{\infty}\sigma_j^{(\text{excl})}} \\ &= \frac{\sigma_{n+1}^{(\text{excl})}\sum_{j=0}^{\infty} R^j_{(n+1)/n}}{\sigma_n^{(\text{excl})} + \sigma_{n+1}^{(\text{excl})}\sum_{j=0}^{\infty} R^j_{(n+1)/n}} \qquad \text{with} \quad R_{(n+1)/n} = \frac{\sigma_{n+1}^{(\text{excl})}}{\sigma_n^{(\text{excl})}} \\ &= \frac{\dfrac{R_{(n+1)/n}\sigma_n^{(\text{excl})}}{1-R_{(n+1)/n}}}{\sigma_n^{(\text{excl})} + \dfrac{R_{(n+1)/n}\sigma_n^{(\text{excl})}}{1-R_{(n+1)/n}}} = \frac{R_{(n+1)/n}}{1-R_{(n+1)/n}+R_{(n+1)/n}} \\ &= R_{(n+1)/n} \; .\end{aligned} \tag{III.11}$$

For the Poisson distribution and the staircase distribution we can summarize the main properties of the n-jet rates in terms of the upper incomplete gamma function $\Gamma(n,\langle n\rangle)$:

	staircase scaling	Poisson scaling
$\sigma_n^{(\text{excl})}$	$\sigma_0^{(\text{excl})}\; e^{-bn}$	$\sigma_0\; \dfrac{e^{-\langle n\rangle}\langle n\rangle^n}{n!}$
$R_{(n+1)/n} = \dfrac{\sigma_{n+1}^{(\text{excl})}}{\sigma_n^{(\text{excl})}}$	e^{-b}	$\dfrac{\langle n\rangle}{n+1}$
$R^{(\text{incl})}_{(n+1)/n} = \dfrac{\sigma_{n+1}}{\sigma_n}$	e^{-b}	$\left(\dfrac{(n+1)\,e^{-\langle n\rangle}\,\langle n\rangle^{-(n+1)}}{\Gamma(n+1)-n\Gamma(n,\langle n\rangle)}+1\right)^{-1}$
$\langle n\rangle$	$\dfrac{1}{2}\dfrac{1}{\cosh b - 1}$	$\langle n\rangle$
P_{pass}	$1-e^{-b}$	$e^{-\langle n\rangle}$

B. Generating functional

A crucial question is if we can derive the Poisson and staircase scaling patterns from QCD. From our discussion of parton splittings we know that the number of radiated jets is well defined after a simple resummation. So-called

generating functionals for the jet multiplicity allow us to calculate resummed jet quantities from first principles in QCD. We construct a generating functional in an arbitrary parameter u by demanding that repeated differentiation at $u = 0$ gives exclusive multiplicity distributions $P_n \equiv \sigma_n/\sigma_{\text{tot}}$,

$$\boxed{\Phi = \sum_{n=1}^{\infty} u^n P_{n-1}} \qquad \text{with} \quad P_{n-1} = \frac{\sigma_{n-1}}{\sigma_{\text{tot}}} = \frac{1}{n!}\frac{d^n}{du^n}\Phi\bigg|_{u=0} . \tag{III.12}$$

For Φ we will suppress the argument u. In the application to gluon emission the explicit factor $1/n!$ corresponds to the phase space factor for identical bosons. Because in P_n we only count radiated jets, our definition uses P_{n-1} where other conventions use P_n. A second observable we can extract from Φ is the average jet multiplicity,

$$\frac{d\Phi}{du}\bigg|_{u=1} = \sum_{n=1}^{\infty} n\, u^{n-1}\, \frac{\sigma_{n-1}}{\sigma_{\text{tot}}}\bigg|_{u=1} = 1 + \frac{1}{\sigma_{\text{tot}}}\sum_{n=1}^{\infty}(n-1)\,\sigma_{n-1} . \tag{III.13}$$

In analogy to the DGLAP equation we can derive an evolution equation for Φ. We start by quoting the fact that for the parton densities and the Sudakov factors defined in Eq.(II.104) the there exists an evolution equation

$$f_i(x,t) = \Delta_i(t,t_0) f_i(x,t_0) + \int_{t_0}^{t}\frac{dt'}{t'}\,\Delta_i(t,t')\sum_j \int_0^{1-\epsilon}\frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,\hat{P}_{i\leftarrow j}(z)\, f_j\left(\frac{x}{z},t'\right) . \tag{III.14}$$

This form reflects our interpretaion of the Sudakov factor as a non-splitting probability. We can motivate the corresponding equation for the generating functional Φ by remembering our probabilistic picture of parton splittings $i \to jk$ in the final state. The splitting particles are then described by generating functionals $\Phi(t)$ instead of parton densities $f(x,t)$, giving us

$$\boxed{\Phi_i(t) = \Delta_i(t,t_0)\Phi_i(t_0) \; + \int_{t_0}^{t}\frac{dt'}{t'}\Delta_i(t,t')\sum_{i\to j,k}\int_0^1 dz\,\frac{\alpha_s}{2\pi}\,\hat{P}_{i\to jk}(z)\,\Phi_j(z^2t')\Phi_k((1-z)^2t')}\,. \tag{III.15}$$

The difference to the DGLAP equation is that the generating functionals just count jets. Unlike for the parton densities the evolution equation does not include a convolution, but instead two generating functionals under the integral. The argument of the strong coupling we assume, without any further motivation, as $\alpha_s(z^2(1-z)^2t')$. It will become clear during our computation that this scale choice is appropriate.

The argument in this section will go two ways: first, we write down a proper differential evolution equation for $\Phi_q(t)$. Then, we solve this equation for quarks, only including the abelian splitting $q \to qg$. To start with, we insert the unregularized splitting kernel from Eq.(II.54) into the evolution equation,

$$\begin{aligned}
\Phi_q(t) &= \Delta_q(t,t_0)\Phi_q(t_0) \; + \int_{t_0}^{t}\frac{dt'}{t'}\Delta_q(t,t')\int_0^1 dz\,\frac{\alpha_s}{2\pi}\,C_F\frac{1+z^2}{1-z}\,\Phi_q(z^2t')\Phi_g((1-z)^2t') \\
&= \Delta_q(t,t_0)\Phi_q(t_0) \; + \int_{t_0}^{t}\frac{dt'}{t'}\Delta_q(t,t')\int_0^1 dz\,\frac{\alpha_s C_F}{2\pi}\,\frac{-(1-z)(1+z)+2}{1-z}\,\Phi_q(z^2t')\Phi_g((1-z)^2t') \\
&= \Delta_q(t,t_0)\Phi_q(t_0) \; + \int_{t_0}^{t}\frac{dt'}{t'}\Delta_q(t,t')\int_0^1 dz\,\frac{\alpha_s C_F}{2\pi}\,\left(\frac{2}{1-z}-1-z\right)\,\Phi_q(z^2t')\Phi_g((1-z)^2t') .
\end{aligned} \tag{III.16}$$

We can simplify the divergent part of Eq.(III.16) using the new integration parameter $t'' = (1-z)^2t'$. This gives us the Jacobian

$$\frac{dt''}{dz} = \frac{d}{dz}(1-z)^2t' = 2(1-z)(-1)t' = -2\frac{t''}{1-z} \qquad \Leftrightarrow \qquad \frac{dz}{1-z} = -\frac{1}{2}\,\frac{dt''}{t''} . \tag{III.17}$$

In addition, we approximate $z \to 1$ wherever possible in the divergent term and cut off all t integrations at the infrared

resolution scale t_0. This will give us the leading logarithm in t space,

$$\int_0^1 dz\ \frac{\alpha_s(z^2(1-z)^2t')C_F}{2\pi}\ \frac{2}{1-z}\ \Phi_q(z^2t')\Phi_g((1-z)^2t') = \Phi_q(t')\int_{t_0}^{t'} dt''\ \frac{\alpha_s(t'')C_F}{2\pi}\frac{1}{t''}\Phi_g(t'')\ . \tag{III.18}$$

For the finite part in Eq.(III.16) we neglect the logarithmic z dependence of all functions and integrate the leading power dependence $1+z$ to 3/2,

$$-\int_0^1 dz\ \frac{\alpha_s(z^2(1-z)^2t')C_F}{2\pi}\ (1+z)\ \Phi_q(z^2t')\Phi_g((1-z)^2t') \simeq -\frac{\alpha_s(t')C_F}{2\pi}\ \frac{3}{2}\ \Phi_q(t')\Phi_g(t')\ . \tag{III.19}$$

Strictly speaking, the strong coupling as well as the two generating functionals could be evaluated at any typical scale covered by the z integral, considering that the prefactor $1+z$ grows towards $z\to 1$; we assume that their change with varying z is small compared to the leading logarithm. After these two simplifying steps Eq.(III.16) reads

$$\begin{aligned}\Phi_q(t) &= \Delta_q(t,t_0)\Phi_q(t_0) + \frac{C_F}{2\pi}\int_{t_0}^t \frac{dt'}{t'}\ \Delta_q(t,t')\left(\int_{t_0}^{t'} dt''\ \frac{\alpha_s(t'')}{t''}\Phi_q(t')\Phi_g(t'') - \frac{3}{2}\alpha_s(t')\Phi_q(t')\Phi_g(t')\right) \qquad \text{(III.20)}\\ &= \Delta_q(t,t_0)\Phi_q(t_0) + \frac{C_F}{2\pi}\Delta_q(t,t_0)\int_{t_0}^t \frac{dt'}{t'}\ \frac{1}{\Delta_q(t',t_0)}\ \Phi_q(t')\left(\int_{t_0}^{t'} dt''\ \frac{\alpha_s(t'')}{t''}\Phi_g(t'') - \frac{3}{2}\alpha_s(t')\Phi_g(t')\right)\ .\end{aligned}$$

The original Sudakov factor $\Delta_q(t,t')$ is split into a ratio of two Sudakov factors. This allows us to differentiate both sides with respect to t,

$$\begin{aligned}\frac{d}{dt}\Phi_q(t) &= \frac{d\Delta_q(t,t_0)}{dt}\Phi_q(t_0) + \frac{C_F}{2\pi}\frac{d\Delta_q(t,t_0)}{dt}\int_{t_0}^t \frac{dt'}{t'}\frac{1}{\Delta_q(t',t_0)}\Phi_q(t')\left(\int_{t_0}^{t'} dt''\frac{\alpha_s(t'')}{t''}\Phi_g(t'') - \frac{3}{2}\alpha_s(t')\Phi_g(t')\right)\\ &\qquad + \frac{C_F}{2\pi}\ \Delta_q(t,t_0)\frac{1}{t}\frac{1}{\Delta_q(t,t_0)}\Phi_q(t)\left(\int_{t_0}^t dt''\ \frac{\alpha_s(t'')}{t''}\Phi_g(t'') - \frac{3}{2}\alpha_s(t)\Phi_g(t)\right)\\ &= \frac{d\Delta_q(t,t_0)}{dt}\left[\Phi_q(t_0) + \frac{C_F}{2\pi}\int_{t_0}^t \frac{dt'}{t'}\frac{1}{\Delta_q(t',t_0)}\Phi_q(t')\left(\int_{t_0}^{t'} dt''\frac{\alpha_s(t'')}{t''}\Phi_g(t'') - \frac{3}{2}\alpha_s(t')\Phi_g(t')\right)\right]\\ &\qquad + \frac{C_F}{2\pi}\ \frac{1}{t}\ \Phi_q(t)\ \left(\int_{t_0}^t dt''\ \frac{\alpha_s(t'')}{t''}\Phi_g(t'') - \frac{3}{2}\alpha_s(t)\Phi_g(t)\right)\\ &= \frac{1}{\Delta_q(t,t_0)}\ \frac{d\Delta_q(t,t_0)}{dt}\ \Phi_q(t) + \Phi_q(t)\frac{C_F}{2\pi}\ \frac{1}{t}\left(\int_{t_0}^t dt''\ \frac{\alpha_s(t'')}{t''}\Phi_g(t') - \frac{3}{2}\alpha_s(t)\Phi_g(t)\right)\ . \qquad \text{(III.21)}\end{aligned}$$

In the last step we use the definition in Eq.(III.20). This simplified equation has a solution which we can write in a closed form, namely

$$\begin{aligned}\Phi_q(t) &= \Phi_q(t_0)\ \Delta_q(t,t_0)\ \exp\left[\frac{C_F}{2\pi}\ \int_{t_0}^t dt'\ \frac{\alpha_s(t')}{t'}\left(\log\frac{t}{t'} - \frac{3}{2}\right)\Phi_g(t')\right]\\ &= \Phi_q(t_0)\ \exp\left[-\int_{t_0}^t dt'\ \Gamma_{q\leftarrow q}(t,t')\right]\ \exp\left[\int_{t_0}^t dt'\ \Gamma_{q\leftarrow q}(t,t')\Phi_g(t')\right]\\ &= \Phi_q(t_0)\ \exp\left[\int_{t_0}^t dt'\ \Gamma_{q\leftarrow q}(t,t')\left(\Phi_g(t')-1\right)\right]\ . \qquad \text{(III.22)}\end{aligned}$$

We can prove this by straightforward differentiation of the first line in Eq.(III.22),

$$\begin{aligned}
\frac{d\Phi_q(t)}{dt} &= \Phi_q(t_0)\,\frac{d\Delta_q(t,t_0)}{dt}\,\exp\left[\frac{C_F}{2\pi}\int_{t_0}^{t}dt'\,\frac{\alpha_s(t')}{t'}\left(\log\frac{t}{t'}-\frac{3}{2}\right)\Phi_g(t')\right] \\
&\quad + \Phi_q(t)\frac{d}{dt}\left[\frac{C_F}{2\pi}\int_{t_0}^{t}dt'\,\frac{\alpha_s(t')}{t'}\left(\log t-\log t'-\frac{3}{2}\right)\Phi_g(t')\right] \\
&= \frac{1}{\Delta_q(t,t_0)}\,\frac{d\Delta_q(t,t_0)}{dt}\,\Phi_q(t)+\Phi_q(t)\,\frac{C_F}{2\pi}\,\frac{\alpha_s(t)}{t}\left(-\log t-\frac{3}{2}\right)\Phi_g(t) \\
&\quad + \Phi_q(t)\,\frac{C_F}{2\pi}\,\frac{1}{t}\int_{t_0}^{t}dt'\,\frac{\alpha_s(t')}{t'}\,\Phi_g(t')+\Phi_q(t)\,\frac{C_F}{2\pi}\,\log t\,\frac{\alpha_s(t)}{t}\,\Phi_g(t) \\
&= \frac{1}{\Delta_q(t,t_0)}\,\frac{d\Delta_q(t,t_0)}{dt}\,\Phi_q(t)+\Phi_q(t)\,\frac{C_F}{2\pi}\,\frac{1}{t}\left(\int_{t_0}^{t}dt'\,\frac{\alpha_s(t')}{t'}\,\Phi_g(t')-\alpha_s(t)\frac{3}{2}\Phi_g(t)\right)\,. \qquad \text{(III.23)}
\end{aligned}$$

The expression given in Eq.(III.22) indeed solves the evolution equation in Eq.(III.21). The corresponding computation for $\Phi_g(t)$ follows the same path.

By definition, the generating functional evaluated at the resolution scale t_0 describes an ensemble of jets which have had no opportunity to split. This means $\Phi_{q,g}(t_0)=u$. The quark and gluon generating functionals to next–to–leading logarithmic accuracy are

$$\begin{aligned}
\Phi_q(t) &= u\,\exp\left[\int_{t_0}^{t}dt'\,\Gamma_{q\leftarrow q}(t,t')\,(\Phi_g(t')-1)\right] \\
\Phi_g(t) &= u\,\exp\left[\int_{t_0}^{t}dt'\left(\Gamma_{g\leftarrow g}(t,t')\,(\Phi_g(t')-1)+\Gamma_{q\leftarrow g}(t')\left(\frac{\Phi_q^2(t')}{\Phi_g(t')}-1\right)\right)\right]\,. \qquad \text{(III.24)}
\end{aligned}$$

The splitting kernels are defined in Eq.(II.106); gluon splitting to quarks described by $\Gamma_{q\leftarrow g}$ is suppressed by a power of the logarithm $\log(t/t')$.

The logarithm $\log(t/t')$ combined with the coupling constant α_s included in the splitting kernels is the small parameter which we will use for the following argument. If this logarithmically enhanced term dominates the physics, the evolution equations for quark and gluons are structurally identical. In both cases, the Φ dependence of the exponent spoils an effective solution of Eq.(III.24). However, the logarithmic form of $\Gamma(t,t')$ ensures that the main contribution to the t' integral comes from the region where $t'\sim t_0$. Unless something drastic happens with the integrands in Eq.(III.24) this means that under the integral we can approximate $\Phi_{q,g}(t_0)=u$ and, if necessary, iteratively insert the solution for $\Phi(t)$ into the differential equation. The leading terms for both, quark and gluon evolution equations turn into the closed form

$$\Phi_{q,g}(t) = u\,\exp\left[\int_{t_0}^{t}dt'\,\Gamma_{q,g}(t,t')\,(u-1)\right] = u\,\exp\left[-(1-u)\int_{t_0}^{t}dt'\,\Gamma_{q,g}(t,t')\right]\,. \qquad \text{(III.25)}$$

Using the Sudakov factor defined in Eq.(II.104) the generating functional in the approximation of large logarithmically enhanced parton splitting is

$$\boxed{\Phi_{q,g}(t) = u\,\Delta_{q,g}(t)^{1-u}}\,. \qquad \text{(III.26)}$$

Its first derivative for general values of u is

$$\begin{aligned}
\frac{d}{du}\,\Phi_{q,g}(t) &= \Delta_{q,g}\frac{d}{du}ue^{-u\log\Delta_{q,g}} \\
&= \Delta_{q,g}\left[e^{-u\log\Delta_{q,g}}+u\,(-\log\Delta_{q,g})\,e^{-u\log\Delta_{q,g}}\right] \\
&= \Delta_{q,g}\,(1-u\log\Delta_{q,g})\,e^{-u\log\Delta_{q,g}}\,. \qquad \text{(III.27)}
\end{aligned}$$

The n-th derivative has the form

$$\frac{1}{n!}\,\frac{d^n}{du^n}\,\Phi_{q,g}(t) = \frac{(-\log\Delta_{q,g})^{n-1}}{n!}\,\Delta_{q,g}\;(n-u\log\Delta_{q,g})\;e^{-u\log\Delta_{q,g}}\;. \tag{III.28}$$

We can show this by induction, starting from the first derivative in Eq.(III.27),

$$\begin{aligned}
\frac{1}{n!}\,\frac{d^n}{du^n}\Phi_{q,g}(t) &= \frac{1}{n}\,\frac{d}{du}\,\left(\frac{1}{(n-1)!}\,\frac{d^{n-1}}{du^{n-1}}\Phi_{q,g}(t)\right)\\
&= \frac{1}{n}\,\frac{d}{du}\left[\frac{(-\log\Delta_{q,g})^{n-2}}{(n-1)!}\,\Delta_{q,g}\;(n-1-u\log\Delta_{q,g})\;e^{-u\log\Delta_{q,g}}\right] \qquad \text{using Eq.(III.28)}\\
&= \frac{(-\log\Delta_{q,g})^{n-2}}{n!}\,\Delta_{q,g}\left[(-\log\Delta_{q,g})\,e^{-u\log\Delta_{q,g}} + (n-1-u\log\Delta_{q,g})\,(-\log\Delta_{q,g})e^{-u\log\Delta_{q,g}}\right]\\
&= \frac{(-\log\Delta_{q,g})^{n-1}}{n!}\,\Delta_{q,g}\;[1+n-1-u\log\Delta_{q,g}]\;e^{-u\log\Delta_{q,g}}\;.
\end{aligned} \tag{III.29}$$

By definition, Eq.(III.28) gives the Poisson scaling pattern in the number of jets, namely

$$P_{n-1} = \Delta_{q,g}(t)\,\frac{|\log\Delta_{q,g}(t)|^{n-1}}{(n-1)!} \qquad \text{or} \qquad \boxed{R_{(n+1)/n} = \frac{|\log\Delta_{q,g}(t)|}{n+1}}\,. \tag{III.30}$$

In addition to the logarithmically enhanced Poisson case we can find a second, recursive solution for the generating functionals. It holds in the limit of small emission probabilities. The emission probability is governed by $\Gamma_{i\leftarrow j}(t,t')$, as defined in Eq.(II.106). We can make it small by avoiding a logarithmic enhancement, corresponding to no large scale ratios t/t_0. In addition, we would like to get rid of $\Gamma_{q\leftarrow g}$ while keeping $\Gamma_{g\leftarrow g}$. Purely theoretically this means removing the gluon splitting into two quarks and limiting ourselves to pure Yang-Mills theory. In that case the scale derivative of Eq.(III.24) reads

$$\begin{aligned}
\frac{d\Phi_g(t)}{dt} &= \Phi_g(t_0)\,\frac{d}{dt}\,\exp\left[\int_{t_0}^t dt'\,\Gamma_{g\leftarrow g}(t,t')\,(\Phi_g(t')-1)\right]\\
&= \Phi_g(t)\,\frac{C_A}{2\pi}\frac{d}{dt}\,\int_{t_0}^t dt'\,\frac{\alpha_s(t')}{t'}\left(\log t-\log t'-\frac{11}{6}\right)\,(\Phi_g(t')-1) \qquad \text{inserting Eq.(II.106)}\\
&= \Phi_g(t)\frac{C_A}{2\pi}\left[\frac{\alpha_s(t)}{t}\left(-\log t-\frac{11}{6}\right)\,(\Phi_g(t)-1)+\frac{d}{dt}\,\log t\,\int_{t_0}^t dt'\,\frac{\alpha_s(t')}{t'}\,(\Phi_g(t')-1)\right]\\
&= \Phi_g(t)\frac{C_A}{2\pi}\left[\frac{\alpha_s(t)}{t}\left(-\log t-\frac{11}{6}\right)\,(\Phi_g(t)-1)+\frac{1}{t}\int_{t_0}^t dt'\,\frac{\alpha_s(t')}{t'}\,(\Phi_g(t')-1)+\log t\,\frac{\alpha_s(t)}{t}\,(\Phi_g(t)-1)\right]\\
&= \Phi_g(t)\frac{C_A}{2\pi t}\left[-\frac{11}{6}\,\alpha_s(t)\,(\Phi_g(t)-1)+\int_{t_0}^t dt'\,\frac{\alpha_s(t')}{t'}\,(\Phi_g(t')-1)\right]\,.
\end{aligned} \tag{III.31}$$

This form is already greatly simplified, but in the combination of the integral and the running strong coupling it is not clear what the limit of small but finite $\log(t/t_0)$ would be. Integrating by parts we find a form which we can estimate systematically,

$$\begin{aligned}
\frac{d\Phi_g(t)}{dt} &= \Phi_g(t)\frac{C_A}{2\pi t}\left[-\frac{11}{6}\alpha_s(t)\,(\Phi_g(t)-1)-\int_{t_0}^t dt'\log t'\frac{d}{dt'}\,(\alpha_s(t')\,(\Phi_g(t')-1))+\log t'\alpha_s(t')\,(\Phi_g(t')-1)\,\Big|_{t_0}^t\right]\\
&= \Phi_g(t)\frac{C_A}{2\pi t}\left[-\frac{11}{6}\alpha_s(t)\,(\Phi_g(t)-1)-\int_{t_0}^t dt'\log\frac{t'}{t_0}\frac{d}{dt'}\,(\alpha_s(t')\,(\Phi_g(t')-1))+\log\frac{t'}{t_0}\alpha_s(t')\,(\Phi_g(t')-1)\,\Big|_{t_0}^t\right]\\
&= \Phi_g(t)\frac{C_A}{2\pi t}\left[\alpha_s(t)\left(\log\frac{t}{t_0}-\frac{11}{6}\right)\,(\Phi_g(t)-1)-\int_{t_0}^t dt'\,\log\frac{t'}{t_0}\,\frac{d}{dt'}\,(\alpha_s(t')\,(\Phi_g(t')-1))\right]\,.
\end{aligned} \tag{III.32}$$

We can evaluate this expression in the limit of $t=t_0+\delta$ or $t_0/t=1-\delta/t$. The two leading terms, ignoring all terms

of the order δ^2, read

$$\begin{aligned}\frac{d\Phi_g(t)}{dt} &= \Phi_g(t)\frac{C_A}{2\pi t}\left[\alpha_s(t)\left(\frac{\delta}{t_0}-\frac{11}{6}\right)(\Phi_g(t)-1)-(t-t_0)\,\frac{\delta}{t}\,\frac{d}{dt}\left(\alpha_s(t)\,(\Phi_g(t)-1)\right)\right]\\ &= \Phi_g(t)\,\frac{C_A}{2\pi}\,\frac{\alpha_s(t)}{t}\left(\frac{\delta}{t_0}-\frac{11}{6}\right)(\Phi_g(t)-1)+\mathcal{O}\left(\delta^2\right)\ . \end{aligned} \tag{III.33}$$

To leading order in δ/t the equation for the generating functional becomes

$$\frac{d\Phi_g(t)}{dt} = \Phi_g(t)\,\tilde\Gamma_{g\leftarrow g}(t,t_0)\ (\Phi_g(t)-1) \qquad \text{with} \qquad \tilde\Gamma_{g\leftarrow g}(t,t_0) = \frac{C_A}{2\pi}\,\frac{\alpha_s(t)}{t}\left(\log\frac{t}{t_0}-\frac{11}{6}\right)\ . \tag{III.34}$$

With $\tilde\Gamma$ we define a slightly modified splitting kernel, where the prefactor α_s/t is evaluated at the first argument t instead of the second argument t_0. Including the boundary condition $\Phi_g(t_0)=u$ we can solve this equation for the generating functional, again using the method of the known solution,

$$\boxed{\Phi_g(t) = \frac{1}{1+\dfrac{1-u}{u\tilde\Delta_g(t)}}} \qquad \text{with} \qquad \tilde\Delta_g(t) = \exp\left(-\int_{t_0}^t dt'\tilde\Gamma_{g\leftarrow g}(t',t_0)\right)\ . \tag{III.35}$$

The derivative of this solution is

$$\begin{aligned}\frac{d\Phi_g(t)}{t} &= \frac{d}{dt}\left(1+\frac{1-u}{u\tilde\Delta_g(t)}\right)^{-1}\\ &= -\Phi_g(t)^2\,\frac{1-u}{u}\,\frac{d}{dt}\exp\left(+\int_{t_0}^t dt'\,\tilde\Gamma_{g\leftarrow g}(t',t_0)\right)\\ &= -\Phi_g(t)^2\,\frac{1-u}{u\tilde\Delta_g(t)}\,\frac{d}{dt}\int_{t_0}^t dt'\,\tilde\Gamma_{g\leftarrow g}(t',t_0)\\ &= -\Phi_g(t)^2\,\left(\frac{1}{\Phi_g(t)}-1\right)\,\frac{d}{dt}\int_{t_0}^t dt'\,\tilde\Gamma_{g\leftarrow g}(t',t_0) = \Phi_g(t)\ (\Phi_g(t)-1)\ \tilde\Gamma_{g\leftarrow g}(t,t_0)\ , \end{aligned} \tag{III.36}$$

which is precisely the evolution equation in Eq.(III.34).

While we have suggestively defined a modified splitting kernel $\tilde\Gamma$ in Eq.(III.34) and even extended this analogy to a Sudakov-like factor in Eq.(III.35) it is not entirely clear what this object represents. In the limit of large $\log(t/t_0)\gg 1$ or $t\gg t_0$, which is not the limit we rely on for the pure Yang–Mills case, we find

$$\begin{aligned}\int_{t_0}^t dt'\,\tilde\Gamma_{g\leftarrow g}(t',t_0) - \int_{t_0}^t dt'\,\Gamma_{g\leftarrow g}(t',t_0) &= \frac{C_A}{2\pi}\,\int_{t_0}^t dt'\,\left(\frac{\alpha_s(t')}{t'}-\frac{\alpha_s(t_0)}{t_0}\right)\,\log\frac{t'}{t_0}\\ &\simeq -\frac{C_A}{2\pi}\,\alpha_s(t_0)\int_{t_0}^t\frac{dt'}{t_0}\,\log\frac{t'}{t_0}\\ &= -\frac{C_A}{2\pi}\,\alpha_s(t_0)\,\left[\frac{t'}{t_0}\log\frac{t'}{t_0}-\frac{t'}{t_0}\right]_1^{t/t_0} = -\frac{C_A}{2\pi}\,\alpha_s(t_0)\,\frac{t}{t_0}\log\frac{t}{t_0}\ . \end{aligned} \tag{III.37}$$

In the staircase limit $t\sim t_0$ and consistently neglecting $\log(t/t_0)$ the two kernels $\Gamma_{g\leftarrow g}$ and $\tilde\Gamma_{g\leftarrow g}$ become identical. In the same limit we find $\Delta_g\sim\tilde\Delta_g\sim 1$. Again using $t'=t_0+\delta$ and only keeping the leading terms in δ we can compute

the leading difference

$$\begin{aligned}
\tilde{\Gamma}_{g\leftarrow g}(t',t_0) - \Gamma_{g\leftarrow g}(t',t_0) &= \frac{C_A}{2\pi}\,\left(\frac{\alpha_s(t')}{t'} - \frac{\alpha_s(t_0)}{t_0}\right)\,\left(\frac{\delta}{t'} - \frac{11}{6}\right) \\
&= -\frac{C_A}{2\pi}\,\frac{11}{6}(t'-t_0)\,\frac{d}{dt}\frac{\alpha_s(t)}{t}\bigg|_{t_0} = -\frac{C_A}{2\pi}\,\frac{11}{6}\delta\,\left[\frac{1}{t}\,\frac{d\alpha_s(t)}{dt} - \frac{\alpha_s(t)}{t^2}\right]_{t_0} \\
&= -\frac{C_A}{2\pi}\,\frac{11}{6}\delta\,\left[-\frac{1}{t}\,\frac{\alpha_s^2(t)b_0}{t} - \frac{\alpha_s(t)}{t^2}\right]_{t_0} \qquad \text{using Eq.(II.22)} \\
&= \frac{C_A\alpha_s(t_0)}{2\pi}\,\frac{11}{6}\,\frac{\delta}{t_0^2}\,\left(1+b_0\alpha_s(t_0)\right) \\
\int_{t_0}^{t} dt'\,\tilde{\Gamma}_{g\leftarrow g}(t',t_0) - \int_{t_0}^{t} dt'\,\Gamma_{g\leftarrow g}(t',t_0) &= \frac{C_A\alpha_s(t_0)}{2\pi}\,\frac{11}{6}\,\frac{\delta^2}{t_0^2}\,\left(1+b_0\alpha_s(t_0)\right)\,. \qquad \text{(III.38)}
\end{aligned}$$

In the pure Yang–Mills theory the running of the strong coupling is described by $b_0 = 1/(4\pi)11N_c/3$. In both limits the true and the modified splitting kernels differ by the respective small parameter.

The closed form for the generating functional in Eq.(III.35) allows us to compute the number of jets in purely gluonic events. The first derivative is

$$\begin{aligned}
\frac{d}{du}\Phi_g(t) &= \frac{d}{du}\,u\left(u+\frac{1-u}{\tilde{\Delta}_g(t)}\right)^{-1} \\
&= \left(u+\frac{1-u}{\tilde{\Delta}_g}\right)^{-1} + u(-1)\left(u+\frac{1-u}{\tilde{\Delta}_g}\right)^{-2}\left(1-\frac{1}{\tilde{\Delta}_g}\right)\,. \qquad \text{(III.39)}
\end{aligned}$$

The form of the n-th derivative we can again prove by induction. Clearly, for $n=1$ the above result is identical with the general solution

$$\frac{d^n}{du^n}\Phi_g(t) = n!\left(\frac{1}{\tilde{\Delta}_g}-1\right)^{n-1}\left(u+\frac{1-u}{\tilde{\Delta}_g}\right)^{-n}\left[1+u\left(u+\frac{1-u}{\tilde{\Delta}_g}\right)^{-1}\left(\frac{1}{\tilde{\Delta}_g}-1\right)\right]\,. \qquad \text{(III.40)}$$

The induction step from $n-1$ to n is

$$\begin{aligned}
\frac{d^n}{du^n}\Phi_g(t) &= (n-1)!\left(\frac{1}{\tilde{\Delta}_g}-1\right)^{n-2}\left\{-(n-1)\left(u+\frac{1-u}{\tilde{\Delta}_g}\right)^{-n}\left(1-\frac{1}{\tilde{\Delta}_g}\right)\left[1+u\,\left(u+\frac{1-u}{\tilde{\Delta}_g}\right)^{-1}\left(\frac{1}{\tilde{\Delta}_g}-1\right)\right]\right. \\
&\qquad \left. + \left(u+\frac{1-u}{\tilde{\Delta}_g}\right)^{-n+1}\left[\left(u+\frac{1-u}{\tilde{\Delta}_g}\right)^{-1}\left(\frac{1}{\tilde{\Delta}_g}-1\right) + u\left(u+\frac{1-u}{\tilde{\Delta}_g}\right)^{-2}\left(\frac{1}{\tilde{\Delta}_g}-1\right)^2\right]\right\} \\
&= (n-1)!\left(\frac{1}{\tilde{\Delta}_g}-1\right)^{n-2}\left\{(n-1)\left(u+\frac{1-u)}{\tilde{\Delta}_g}\right)^{-n}\left(\frac{1}{\tilde{\Delta}_g}-1\right)\left[1+u\,\left(u+\frac{1-u}{\tilde{\Delta}_g}\right)^{-1}\left(\frac{1}{\tilde{\Delta}_g}-1\right)\right]\right. \\
&\qquad \left. + \left(u+\frac{1-u}{\tilde{\Delta}_g}\right)^{-n}\left(\frac{1}{\tilde{\Delta}_g}-1\right)\left[1+u\left(u+\frac{1-u}{\tilde{\Delta}_g}\right)^{-1}\left(\frac{1}{\tilde{\Delta}_g}-1\right)\right]\right\} \\
&= n!\left(\frac{1}{\tilde{\Delta}_g}-1\right)^{n-1}\left(u+\frac{1-u}{\tilde{\Delta}_g}\right)^{-n}\left[1+u\left(u+\frac{1-u}{\tilde{\Delta}_g}\right)^{-1}\left(\frac{1}{\tilde{\Delta}_g}-1\right)\right]\,. \qquad \text{(III.41)}
\end{aligned}$$

Evaluating Eq.(III.40) for $u = 0$ gives us the jet rates

$$P_{n-1} = \frac{1}{n!}\frac{d^n}{du^n}\Phi_g(t)\bigg|_{u=0} = \left(\frac{1}{\tilde{\Delta}_g} - 1\right)^{n-1} \tilde{\Delta}_g^n = \tilde{\Delta}_g \left(1 - \tilde{\Delta}_g\right)^{n-1} , \tag{III.42}$$

which predicts constant ratios

$$\boxed{R_{(n+1)/n} = 1 - \tilde{\Delta}_g(t)} . \tag{III.43}$$

Such constant ratios define a staircase pattern. It has for a long time been considered an accidental sweet spot where many QCD effects cancel each other to produce constant ratios of successive exclusive n-jet rates. Our derivation from the generating functionals suggest that staircase scaling is one of two pure jet scaling patterns:

1. in the presence of large scale differences abelian splittings generate a Poisson pattern with $R_{(n+1)/n} \propto 1/(n+1)$, as seen in Eq.(III.30).

2. for democratic scales non–abelian splittings generate a staircase pattern with constant $R_{(n+1)/n}$ shown in Eq.(III.43).

C. Jets finders

Until now we have consistently pretended that quarks and gluons produced at the LHC are what we observe in the detectors. In perturbative QCD they are assumed to form the initial and final states, even though they cannot exist individually as long as QCD is asymptotically free. Indeed, the gluon and all quarks except for the top quark hadronize before they decay and form bunches of baryons and mesons which in turn decay in many stages. At the LHC these particles carry a lot of energy, typically around the electroweak scale. Relativistic kinematics then tells us that these baryons and mesons are strongly boosted together to form jets. Those jets is what we measure at hadron colliders and what we have to link to the partons produced in the hard interaction.

Imagine we observe a large number of energy depositions in the ATLAS or CMS calorimeter which we would like to combine into jets. We know that they come from a small number of partons which originate in the hard QCD process and which since have undergone a sizeable number of splittings, hadronized and decayed to stable particles. Can we try to reconstruct the original partons?

We know our previous discussions that in QCD we can produce an arbitrary number of hard jets in a hard matrix element and another arbitrary number of jets via soft or collinear radiation. A major theoretical issue is that we can radiate an arbitrary number of soft partons from the hard process without affecting the hard process, as described by the exponentiation of soft radiation in the eikonal approximation. Any jet algorithm has to respect this feature. Similarly, any jet algorithm needs to include collinear radiation in a way that the limit of two collinear constituents to one combined constituent is smooth. In this case the reconstruction of the partons through the jet algorithm is infrared and collinear safe.

We illustrate the behavior of possible jet algorithms facing soft radiation off a hadronically decaying W-boson in Fig. 5. At tree level, any jet algorithm will typically find two well-separated jets, corresponding to two hard partons

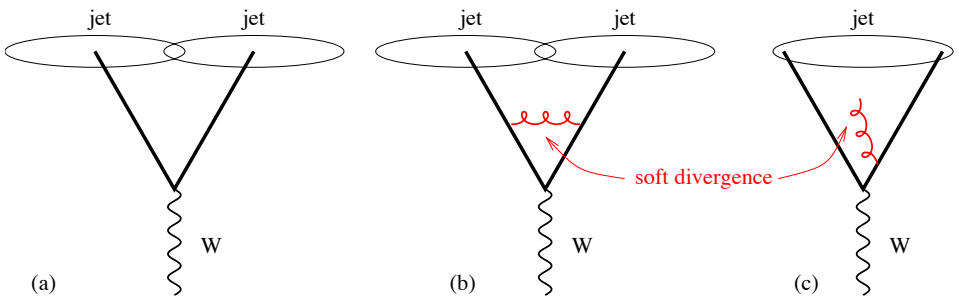


FIG. 5: Parton configurations interpreted through a collinear safe (left) and collinear unsafe (right) algorithm. Figure and example from Mike Seymour.

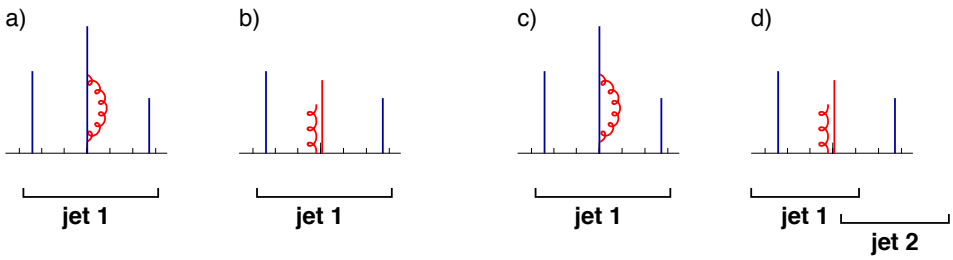


FIG. 6: Parton configurations interpreted through a collinear safe (left) and collinear unsafe (right) algorithm. Figure and example from Gavin Salam.

possibly accompanied by soft and collinear radiation. To next-to-leading order in perturbation theory we can add soft gluons, both virtual and real. The two infrared divergences cancel, as long as the jet algorithm always combines the two contribution into the same final state configuration or hard process. Let us assume a naive cone algorithm, simply combining jet constituents based on their geometric separation. In this case the leading order configuration (a) and the virtual correction (b) will be treated exactly the same way. In both cases there is no way to combine the two widely separated constituents into one jet, so the algorithm identifies two hard partons in the final state. In contrast, when the soft gluon radiated between the two quarks (c) serves as the starting point of the jet algorithm, the algorithm will be able to combine both hard quarks with the central gluon into one jet. A as result, the virtual correction contributes to a two-parton configuration, while the real correction contributes to a one-parton configuration, and the infrared divergences will not cancel. An infrared save algorithm has to recognize that a soft gluon cannot serve as a starting point for the parton reconstruction. Instead, it has to start from one of the harder jets and construct two partons. To which of the two its adds the soft gluon does not matter for the infrared safety.

Collinear safety is illustrated in Fig. 6. We assume a configuration with three hard partons, where the hardest parton can either include a virtual gluon contribution (a,c) or undergo a collinear splitting (b,d). Again, we know that the virtual and real emission diagrams lead to individual divergences, and only their combination gives us a finite result. This means that the jet algorithm has to combine the constituents the same way, as illustrated in (a,b). A problem occurs for example when a cone algorithm starts from the hardest constituent, combined with a distance measure. For the collinear splitting it would start from the hardest constituents to the left and add the central constituent. However, the right constituent would be too far away and not be included. As before, the algorithm splits the two divergent contributions and leads to individually divergent rates for the one-parton and two-parton rates.

From the above discussion we learn that jet algorithms starting from a fixed seed of constituents and then combining them based on a geometric distance measure will destroy the features of quantum field theory on the level of reconstructed partons. To avoid this, we define jets is based on so-called recombination algorithms, including for example the Cambridge–Aachen or (anti-) k_T algorithms.

The basic idea of recombination algorithms is to ask if a given subjet has a soft or collinear partner. We know that partons produced in a hard process preferably turn into collinear pairs of partons as approximately described by the parton shower. To decide if two subjets have arisen from one parton leaving the hard process we have to define a collinearity measure. This measure will on the one hand include the distance in R space, where $(\Delta R)^2 = (\Delta\eta)^2 + (\Delta\phi)^2$, and on the other hand the transverse momentum of one subjet with respect to another or to the beam axis. Explicit measures weighted by the relative power of the two ingredients are

$$\begin{aligned}
k_T \qquad & y_{ij} = \frac{\Delta R_{ij}}{R} \, \min\left(p_{T,i}, p_{T,j}\right) \qquad & y_{iB} = p_{T,i} \\
\text{C/A} \qquad & y_{ij} = \frac{\Delta R_{ij}}{R} & y_{iB} = 1 \\
\text{anti-}k_T \qquad & y_{ij} = \frac{\Delta R_{ij}}{R} \, \min\left(p_{T,i}^{-1}, p_{T,j}^{-1}\right) & y_{iB} = p_{T,i}^{-1} \, .
\end{aligned} \tag{III.44}$$

The parameter R balances the jet–jet and jet–beam criteria.

In an exclusive jet algorithm we define two subjets as coming from one jet if $y_{ij} < y_{\text{cut}}$, where y_{cut} is a reference scale we give to the algorithm. The reason to introduce this parameter is that quantum field theory tells us that hard partons can split into any number of soft and collinear partons. From first principles we do not know how many constituents it should combine to a hard parton. Therefore, we need to tell the jet algorithm either how many jets it

should arrive at or what the resolution of the smallest subjets we consider partons should be, whatever the measure for this resolution might be. An exclusive jet algorithm then proceeds as

(1) for all combinations of two subjets in the event find the minimum $y^{\min} = \min_{ij}(y_{ij}, y_{iB})$

(2a) if $y^{\min} = y_{ij} < y_{\text{cut}}$ merge subjets i and j and their momenta, keep only the new subjet i, go back to (1)

(2b) if $y^{\min} = y_{iB} < y_{\text{cut}}$ remove subjet i, call it beam radiation, go back to (1)

(2c) if $y^{\min} > y_{\text{cut}}$ keep all subjets, call them jets, done

The result of the algorithm will of course depend on the resolution y_{cut}. Alternatively, we can give the algorithm the minimum number of physical jets and stop there. In an inclusive jet algorithm we avoid introducing y_{cut}. Instead, y_{iB} acts as the universal cutoff and sets the scale for all jet–jet separations.

In the C/A example we immediately see that this translates into a geometric jet size given by R. For regular QCD jets we choose values of $R = 0.4...0.7$. For the C/A and k_T cases we see that an inclusive jet algorithm produces jets arbitrarily close to the beam axis. Those are hard to observe and often not theoretically well defined, as we know from our discussion of collinear divergences. Therefore, inclusive jet algorithms typically include a minimum cut on $p_{T,\text{jet}}$ which at the LHC can be anything from 20 GeV to more than 100 GeV, depending on the analysis.

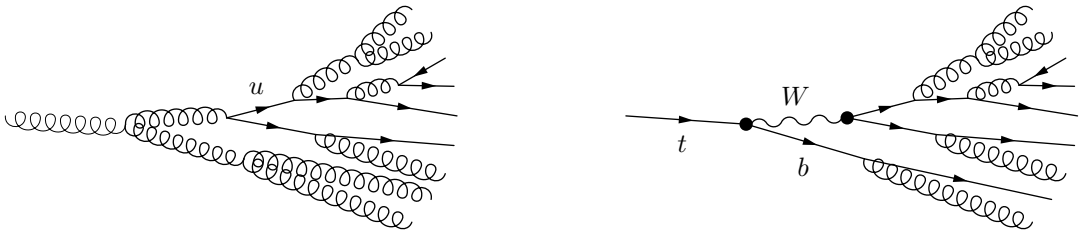
A technical question is what 'combine jets' means in terms of the four-momentum of the new jet. The three-momentum vectors we simply add $\vec{k}_i + \vec{k}_j \to \vec{k}_i$. For the zero component we can assume that the new physical jet have zero invariant mass, which is inspired by the massless parton we are usually looking for. If instead we add the four-momenta we can compute the invariant mass of the jet constituents, the jet mass. As we will see in the next section this allows us to extend the concept of jet algorithms to massive particles like a W or Z boson, the Higgs boson, or the top quark.

All jet algorithms them have in common that they link physical objects, namely calorimeter towers, to other more or less physical objects, namely partons from the hard process. As we can see from the different choices in Eq.(III.44) we have all the freedom in the world to weight the angular and transverse momentum distances relative to each other. As determined by their power dependence on the transverse momenta, the three algorithms start with soft constituents (k_T), purely geometric (Cambridge–Aachen C/A) or hard constituents (anti-k_T) to form a jet. While for the k_T and the C/A algorithms it is fairly clear that the intermediate steps have a physical interpretation, this is not clear at all for the anti-k_T algorithm.

The problem of the theoretically preferred k_T algorithm arises with pile–up or underlying event, *i.e.* very soft QCD activity entering the detectors undirectionally or from secondary partonic vertices. Such noise is easiest understood geometrically in a probabilistic picture. Basically, the low energy jet activity is constant all over the detector, so we subtract it from each event. How much energy deposition we have to subtract from a reconstructed jet depends on the area the jet covers in the detector. Therefore, it is a major step that even for the k_T algorithm we can compute an IR–safe geometric jet size. The C/A and anti-k_T algorithms are more compact and easier to interpret experimentally.

D. Fat jets

Starting from the way the experiments at the Tevatron and the LHC search for bottom jets, including several detailed requirements on the content of such jets, the question arises if we can in general look for other heavy objects inside a jet. Such jets involving heavy particles and (usually) a large geometrical size are referred to as fat jets. For example, looking for boosted top quarks a fat jet algorithm will try to distinguish between two splitting histories, where we mark the massive splittings from boosted top decays:



The splittings inside the light–flavor jet are predicted by the soft and collinear QCD structure. The splittings in the top decays differ because some of the particles involved have masses. This is the jet substructure pattern a fat jet algorithm looks for.

Three main motivations lead us into the direction of fat jets: first, dependent on our physics model heavy objects like W bosons or top quarks will be boosted enough to fit into a regular jet of angular size $R \lesssim 0.7$. Secondly, a jet algorithm might include hadronic decay products which we would not trust to include in a regular mass reconstruction based on reconstructed detector objects. And finally, even if only a fraction of the heavy particles we are searching for are sufficiently boosted such an algorithm automatically resolves signal combinatorics known to limit some LHC analyses.

At the LHC, we are guaranteed to encounter the experimental situation $p_T/m \gtrsim 1$ for electroweak gauge bosons, Higgs bosons, and top quarks. The more extreme case of $p_T \gg m$, for example searching for top quarks with a transverse momentum in excess of 1 TeV, is unlikely to appear in the Standard Model and will only become interesting if we encounter very heavy resonances decaying to a pair of top quarks. This is why we focus on the moderate scenario. Amusingly, the identification of W and top jets was part of the original paper studying the pattern of splittings y_{ij} defining the k_T algorithm. At the time this was mostly a gedankenexperiment to test the consistency of the general k_T algorithm approach. Only later reality caught up with it.

Historically, fat jet searches were first designed to look for strongly interacting W bosons. Based on the k_T algorithm they look for structures in the chain of y values introduced in Eq.(III.44), which define the kinematics of each jet. For such an analysis of y values it is helpful but not crucial that the intermediate steps of the jet algorithm have a physics interpretation. More recent fat jet algorithms looking for not too highly boosted heavy particles are based on the C/A algorithm which appears to be best suited to extract massive splittings inside the jet clustering history. A comparison of different jet algorithms can be found in the original paper on associated Higgs and gauge boson production. Using a C/A algorithm we can search for hadronically decaying boosted W and Z bosons. The problem is that for those we only have one hard criterion based on which we can reject QCD backgrounds: the mass of the W/Z resonance. Adding a second W/Z boson and possibly the mass of a resonance decaying to these two, like a heavy Higgs boson, adds to the necessary QCD rejection. For Higgs and top decays the situation is significantly more promising.

Starting with the Higgs tagger we search for jets which include two bottom quarks coming from a massive Higgs boson. First, we run the C/A algorithm over the event, choosing a large geometric size $R = 1.2$ estimated to cover

$$R_{b\bar{b}} \sim \frac{1}{\sqrt{z(1-z)}} \frac{m_H}{p_{T,H}} > \frac{2m_H}{p_{T,H}} \; , \tag{III.45}$$

in terms of the transverse momentum of the boosted Higgs and the momentum fractions z and $1-z$ of the two bottom jets. We then analyzes individual clusterings where the original jet algorithm combines j_1 and j_2 into j:

1. Un-do the last step of the fat jet clustering where the (parent) subjet j breaks into two (daughter) subjets $j_{1,2}$. For the Higgs tagger we use the C/A jet algorithm.

2. Test three conditions: first, the drop in jet mass has to be large for a heavy particle decay; second, the splitting should then be symmetric; finally the subjets $j_{1.2}$ have to be sufficiently hard:

$$\frac{\min m_{j_i}}{m_j} < 0.67 \qquad \frac{\min(p_{T,j_i}^2)\Delta R_{j_1 j_2}^2}{m_j^2} \sim \frac{\min p_{T,j_i}}{\max p_{T,j_i}} > 0.09 \qquad p_{T,j_i} > 30 \text{ GeV} \; , \tag{III.46}$$

3. For all other splittings, identify the more massive subjet of the $j_{1.2}$ with j and remove the less massive one. This splitting is then removed from the relevant splitting history of the fat jet.

4. Go to the next splittings with the parent subjet $j \equiv j_1$ and, if applicable, $j \equiv j_2$. The un-clustering loop will stop once the last condition in Eq.(III.46) cannot be met anymore.

5. Reconstruct the Higgs mass from the jet mass m_j of the relevant splitting(s) and add b-tagging information.

Applying jet algorithms with very large R size makes us increasingly vulnerable to underlying event, pile–up, or even regular initial–state radiation. This means that we cannot simply use the mass of a set of fat jet constituents. Instead, we apply a filtering procedure looking at the same constituent with a higher resolution which can for example be $R_{\text{filt}} = \min(0.3, R_{b\bar{b}}/2)$. This filtering serves several purposes: first, it significantly reduces the y-ϕ surface area of

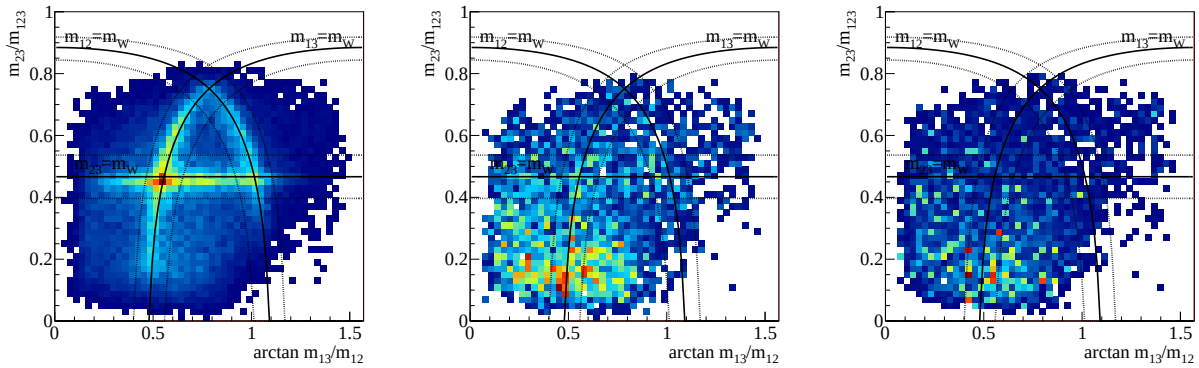


FIG. 7: Events in the arctan m_{13}/m_{12} vs m_{23}/m_{123} plane for $t\bar{t}$ (left). W+jets (center) and pure QCD jets (right) samples. More densely populated regions of the phase space appear in red.

the relevant constituents and thereby the destructive power of the underlying event and pile–up. Second, it allows us to explicitly include gluon radiation off the $b\bar{b}$ pair to improve the Higgs mass resolution by including the three hardest filtered constituents, two bottom quarks and a radiated gluon.

This way, the Higgs taggers relies only on one kinematic criterion, the mass of the $b\bar{b}$ pair, combined with two bottom tags of the constituents The combined QCD rejection is at least 10^{-4}, which might even allows us to run a Higgs tagger over any kind of event sample and see if we find any Higgs bosons for example in new physics decays.

Interestingly, the most frequently used fat jet taggers target hadronically decaying top quarks. They start from a C/A jet of size $R = 1.5 - 1.8$ and search for mass drops corresponding to the top and W masses. For example the original HEPTopTagger algorithm proceeds in steps similar to the Higgs tagger:

1. Un-do the last clustering of the jet j; apply the mass drop criterion $\min m_{j_i} < 0.8\, m_j$ to determine if we keep j_1 and j_2; ignore subjets with $m_{j_i} < 30$ GeV to eventually end the un-clustering.

2. Apply a filtering stage with up to five relevant constituents to construct one three-subjet combination with a jet mass within $m_t \pm 25$ GeV. This defines a sphere in the space of pair-wise invariant masses, illustrated in Fig. 7.

3. Order these three subjets by p_T. If their jet masses (m_{12}, m_{13}, m_{23}) satisfy one of the following three criteria, accept them as a top candidate:

$$
\begin{aligned}
&0.2 < \arctan\frac{m_{13}}{m_{12}} < 1.3 \qquad \text{and} \quad R_{\text{min}} < \frac{m_{23}}{m_{123}} < R_{\text{max}} \\
&R^2_{\text{min}}\left(1+\left(\frac{m_{13}}{m_{12}}\right)^2\right) < 1-\left(\frac{m_{23}}{m_{123}}\right)^2 < R^2_{\text{max}}\left(1+\left(\frac{m_{13}}{m_{12}}\right)^2\right) \quad \text{and} \quad \frac{m_{23}}{m_{123}} > R_{\text{soft}} \\
&R^2_{\text{min}}\left(1+\left(\frac{m_{12}}{m_{13}}\right)^2\right) < 1-\left(\frac{m_{23}}{m_{123}}\right)^2 < R^2_{\text{max}}\left(1+\left(\frac{m_{12}}{m_{13}}\right)^2\right) \quad \text{and} \quad \frac{m_{23}}{m_{123}} > R_{\text{soft}}
\end{aligned}
\tag{III.47}
$$

4. For consistency, require the combined p_T of the three subjets to be above 200 GeV.

The dimensionless mass windows $R_{\text{min}} = 85\% \times m_W/m_t$ and $R_{\text{max}} = 115\% \times m_W/m_t$ are tunable. The soft cutoff $R_{\text{soft}} = 0.35$ removes QCD events which the C/A algorithm cannot correctly identify as soft radiation.

Kinematically, the interesting feature in Fig. 7 is that for the top signal two invariant masses tend to correspond to m_W. This is related to the endpoint of the larger of the two invariant masses from the b-jet and the light-flavor jets in top decays,

$$
m_{bj}^{\text{max}} = \sqrt{m_t^2 - m_W^2} \approx 155 \text{ GeV} \,. \tag{III.48}
$$

This means that one of the two combinations will typically give $m_{bj} \approx m_W$ by accident, and it makes no sense to ask the algorithm to identify the correct W-decay jets. Hadronic top tagging is by now standard in ATLAS and CMS and allows us to search for new physics just using hadronically decaying top quarks. The general field of jet

substructure is a great success story at the LHC and arguably the most innovative fields in LHC physics, both in theory and experiment.