Stefano Frixione

QCD: are we ready for the LHC?

Academic training – Lecture 2

CERN, 4 – 7/12/2006

Summary of lecture 1

- Understanding QCD is crucial for the LHC discovery programme
- QCD is an asymptotically-free QFT, supported by hadron spectroscopy and high-energy experiments
- Perturbative techniques can be used, but are not sufficient: long-distance effects are always present
- To deal with them, one must introduce (at least) hadron-parton duality, infrared safety, factorization theorems

EXAMPLES OF PERTURBATION THEORY AT WORK

$e^+e^- \longrightarrow$ hadrons, jets

DIS and the problem of initial-state divergences

Scale dependence of PDFs

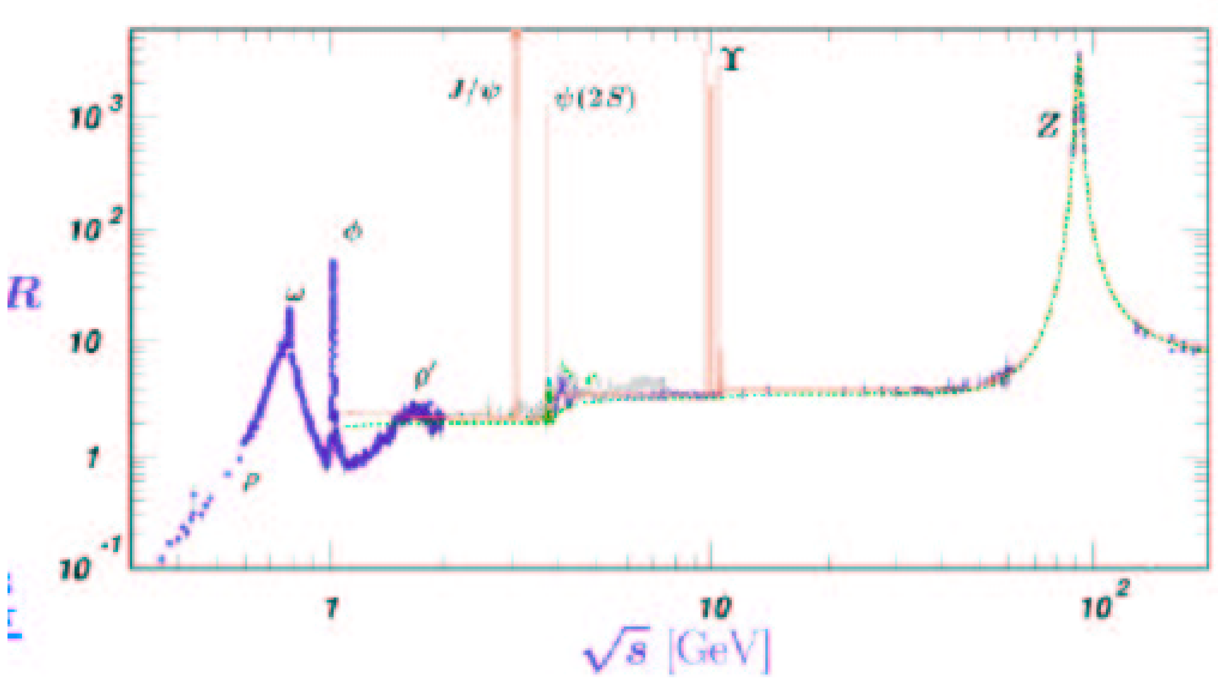
Let's see in practice the way in which hadron-parton duality, infrared safety and factorization theorems work

The simplest case is the total hadronic rate in e^+e^- collisions

- Hadron-parton duality $\Longrightarrow$ compute the total *partonic* rate
- Total rate is (trivially) infrared safe

It's actually customary to give the results as

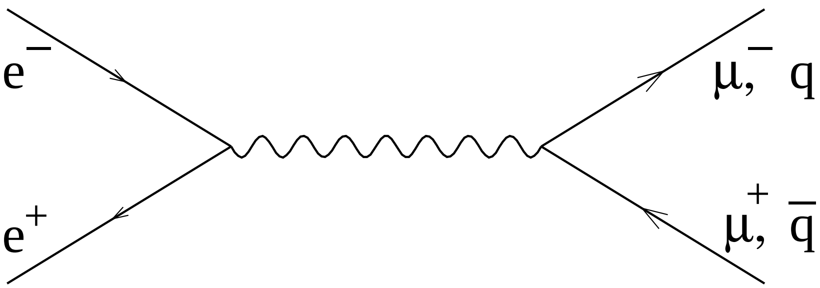
$$R = \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)}$$



At the lowest order in perturbation theory (of α_S)

$$R = \sum_{i,f} \frac{\sigma(e^+e^- \to q_i^{(f)}\bar{q}_i^{(f)})}{\sigma(e^+e^- \to \mu^+\mu^-)} \longrightarrow \sum_{i=1}^{N_C}\sum_{f=1}^{N_F} e^2(f)$$

since numerator and denominator are the same diagram



$$\Longrightarrow \quad R = N_C\left[\left(\frac{2}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 + \left(-\frac{1}{3}\right)^2\right]$$

which is also a test of colour and charge assignments (here for u, d, and s quarks)

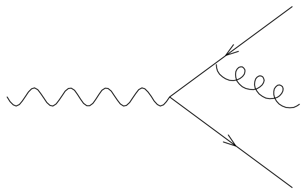
For larger c.m. energies, just add more quark flavours

- Is this result systematically improvable, in the sense of perturbation theory? This is what we expect from the β_{QCD} computation

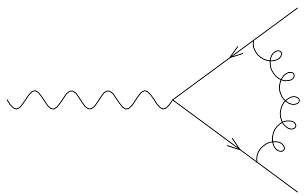
Perturbative corrections to R

At the first order beyond Born (*next-to-leading order*, NLO), there are two classes of corrections (as in QED)

- Real contribution: all Feynman diagrams with an additional (wrt Born) parton in the final state



- Virtual contribution: all one-loop Feynman diagrams that can be obtained from Born diagrams

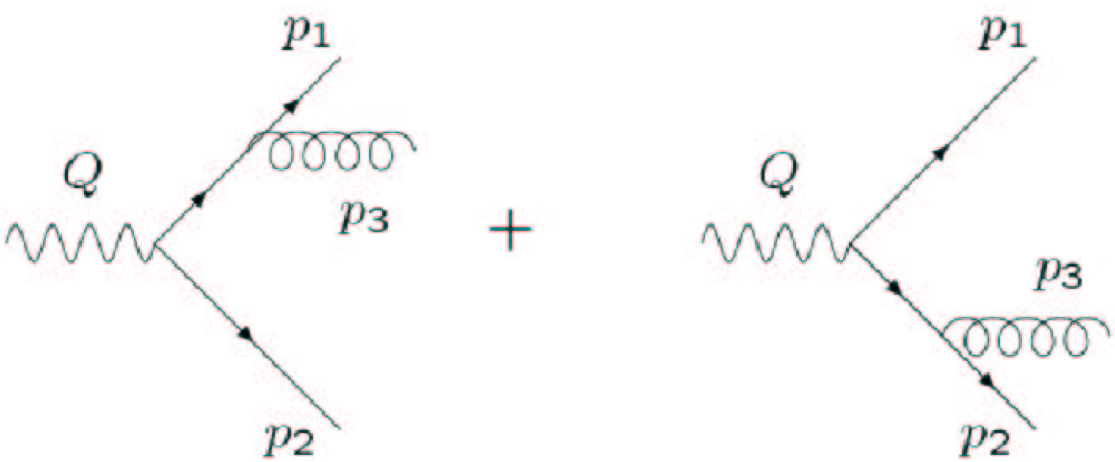


R and V don't interfere: diagrams have different number of legs

$$\text{real} = g_S \mathcal{A}_R \qquad \text{virtual} = g_S^2 \mathcal{A}_V$$

$$|\mathcal{A}_{NLO}|^2 = |\mathcal{A}_{LO}|^2 + \alpha_S \left(|\mathcal{A}_R|^2 + 2\Re(\mathcal{A}_{LO}\mathcal{A}_V^\star) \right) + \mathcal{O}(\alpha_S^2)$$

Real contribution



$$x_i = \frac{2p_i \cdot Q}{Q^2} = \frac{2E_i}{\sqrt{s}}$$

$$p_1 + p_2 + p_3 = Q \implies$$

$$x_1 + x_2 + x_3 = 2$$

Phase space and matrix element:

$$d\Phi_{q\bar{q}g} = \frac{s}{32(2\pi)^5}\delta(2 - x_1 - x_2 - x_3)dx_1 dx_2 dx_3 d\Omega$$

$$|\mathcal{A}_R|^2 = |\mathcal{A}_{LO}|^2 \, C_F \frac{\alpha_S}{2\pi} \frac{x_1^2 + x_2^2}{(1-x_1)(1-x_2)}$$

which lead to

$$\sigma_R = \int d\Phi_{q\bar{q}g} \, |\mathcal{A}_R|^2 = \infty$$

It is instructive to see why this is divergent

$$
\begin{aligned}
1 - x_1 &= x_2 \frac{E_3}{\sqrt{s}}(1 - \cos\theta_{23}) = \frac{(p_2 + p_3)^2}{Q^2} \\
1 - x_2 &= x_1 \frac{E_3}{\sqrt{s}}(1 - \cos\theta_{13}) = \frac{(p_1 + p_3)^2}{Q^2}
\end{aligned}
$$

The divergences of the matrix elements are at

$$
\begin{aligned}
x_1 \longrightarrow 1 \;\&\; x_2 \longrightarrow 1 \quad &\Longleftrightarrow \quad E_3 \longrightarrow 0 \qquad \text{soft} \\
x_1 \longrightarrow 1 \quad &\Longleftrightarrow \quad \theta_{23} \longrightarrow 0 \qquad \text{collinear} \\
x_2 \longrightarrow 1 \quad &\Longleftrightarrow \quad \theta_{13} \longrightarrow 0 \qquad \text{collinear}
\end{aligned}
$$

This clarifies that the divergences are not physical: we are pushing pQCD beyond its range of applicability, since parton energies or parton-pair invariant masses are comparable to hadron masses $\Longrightarrow$ confinement effects can't be neglected

In other words: we are trying to resolve partons in a regime where the concept of parton is not particularly meaningful

- Go home and throw hadron-parton duality (and pQCD) in the bin?

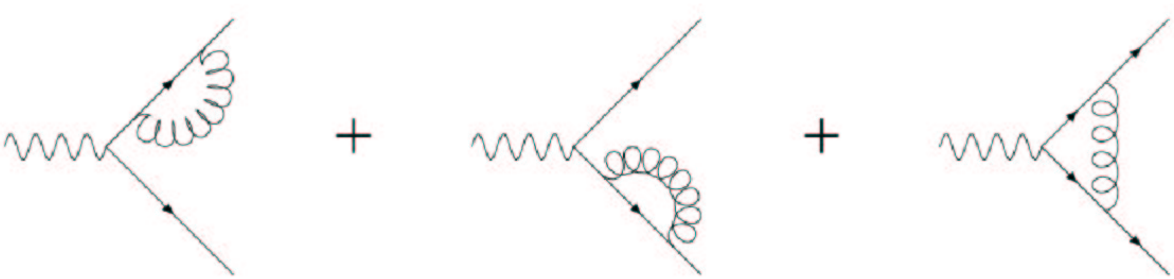
Not yet: what the previous computation tells us is that the cross section for the production of $q\bar{q}g$ is not a meaningful quantity in perturbation theory

But this cross section is just one of the contributions to $e^+e^- \longrightarrow$ hadrons at $\mathcal{O}(\alpha_S)$ – we still have to consider the virtual contribution

So before throwing everything away, we have to prove that soft/collinear emissions are dominant also after adding virtual corrections

Note that what we've got is not peculiar of QCD: you get the same if you compute $\mu^+\mu^-\gamma$ production in QED

Virtual contribution



$$x_i = \frac{2p_i \cdot Q}{Q^2} = \frac{2E_i}{\sqrt{s}}$$
$$p_1 + p_2 = Q \Longrightarrow$$
$$x_1 = 1, \ x_2 = 1$$

One can easily see that

$$\sigma_V = \int d\Phi_{q\bar{q}} \Re(\mathcal{A}_{LO} \mathcal{A}_V^\star) = -\infty$$

- Physical meaning: we are trying to compute the probability of having *exactly* two quarks in the final state
- As in QED, this quantity diverges order-by-order in PT. The result to all orders, however, is not the same as in QED, owing to the different behaviour of the running coupling

$$\sigma_R + \sigma_V = \infty - \infty = ?$$

! *Regularize* R and V contributions before summing them $\longrightarrow$ in QCD, this usually means computing the integrals in $d = 4 - 2\epsilon$ dimensions

$$\int^1 \frac{dx}{1-x} = -\log(0) \xrightarrow{\text{regularization}} \int^1 \frac{dx(1-x)^{-2\epsilon}}{1-x} = -\frac{1}{2\epsilon}$$

$$\sigma_R = \sigma_{LO} C_F \frac{\alpha_S}{2\pi} \left(\frac{2}{\epsilon^2} + \frac{3}{\epsilon} + \frac{19}{2} - \pi^2 \right) + \mathcal{O}(\epsilon)$$

$$\Longrightarrow \quad \sigma_V = \sigma_{LO} C_F \frac{\alpha_S}{2\pi} \left(-\frac{2}{\epsilon^2} - \frac{3}{\epsilon} - 8 + \pi^2 \right) + \mathcal{O}(\epsilon)$$

$$\lim_{\epsilon \to 0} (\sigma_R + \sigma_V) = \frac{\alpha_S}{\pi} \sigma_{LO}$$

The singularities are gone! So we can obtain

$$R = N_C \sum_f Q_f^2 \left(1 + \frac{\alpha_S}{\pi} \right) + \mathcal{O}(\alpha_S^2)$$

This is a small correction ($< 5\%$), and improves the comparison to data – we have proven that the total rate is insensitive to soft/collinear emissions

Physical meaning: soft/collinear real configurations are kinematically degenerate with virtual configurations. Thus, it looks like finite quantities are obtained by summing over degenerate (ie non-resolvable) partonic configurations

This is true to all orders:

Kinoshita-Lee-Nauenberg (KLN) theorem: in the computation of inclusive (enough) quantities, infrared divergences cancel, and the result is finite

And this can indeed be checked by explicit computations $\longrightarrow$

$$R = R_{LO}\left[1 + \frac{\alpha_S}{\pi} + 1.411\left(\frac{\alpha_S}{\pi}\right)^2 - 12.8\left(\frac{\alpha_S}{\pi}\right)^3\right] + \mathcal{O}(\alpha_S^4)$$

The new terms improve further the agreement with data

This is a huge success! Keep in mind we have used several highly non trivial ingredients

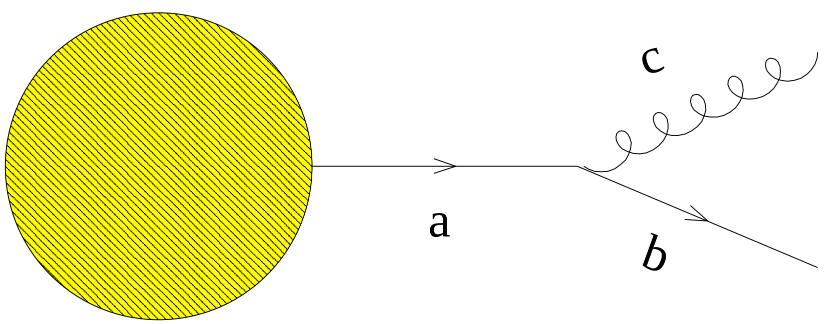
- *Asymptotic freedom*
- *Hadron-parton duality*
- *Infrared safety*

and we have also verified that the KLN theorem works

Speaking of which: how can one prove such an all-order statement?

How can one prove such a general statement as KLN theorem, since computations are observable-specific?

Divergences are actually observable independent, and "universal"; can be easily computed in a physical gauge



$$k_b = z k_a + k_T + \zeta_b n$$

$$k_c = (1-z) k_a - k_T + \zeta_c n$$

$$k_b^2 = 0 \;\Rightarrow\; \zeta_b = -\frac{k_T^2}{2 z n \cdot k_a}$$

$$k_c^2 = 0 \;\Rightarrow\; \zeta_c = -\frac{k_T^2}{2(1-z) n \cdot k_a}$$

$$d\sigma_R = \frac{\alpha_S}{2\pi} \int dk_T^2 dz\, C_F \frac{1+z^2}{1-z} \frac{1}{k_T^2} d\sigma^{(0)}(k_a) + \text{non singular}$$

Again the collinear ($k_T \to 0$) and soft ($z \to 1$, with $k_T \to (1-z)\hat{k}_T$) divergences. They arise when parton a goes on shell $\Longrightarrow$ the propagator diverges

These IR divergences will cancel when adding virtual corrections

The quantity associated with the divergence depends only on parton flavours and kinematics. At the LO, we have the following cases

$$q \to q(z)g(1-z) \quad \Longrightarrow \quad P_{qq}(z) = C_F \frac{1+z^2}{1-z}$$

$$g \to q(z)\bar{q}(1-z) \quad \Longrightarrow \quad P_{qg}(z) = T_R \left(z^2 + (1-z)^2\right)$$

$$q \to g(z)q(1-z) \quad \Longrightarrow \quad P_{gq}(z) = C_F \frac{1+(1-z)^2}{z} = P_{qq}(1-z)$$

$$g \to g(z)g(1-z) \quad \Longrightarrow \quad P_{gg}(z) = C_A \left(\frac{z}{1-z} + \frac{1-z}{z} + z(1-z)\right)$$

$$C_F = \frac{4}{3}\,, \quad C_A = 3\,, \quad T_R = \frac{1}{2}$$

which are the (unsubtracted) *Altarelli-Parisi splitting kernels*

In summary

- When considering perturbative corrections, IR divergences appear
- The residues of the IR divergences are independent of the production process, and can be easily computed
- Certain observables are finite, ie insensitive to the IR sector. For this to happen, real and virtual contributions to the perturbative corrections must both be considered at the NLO
- Perturbative corrections are larger than in QED, but still under control; a pQCD program makes sense

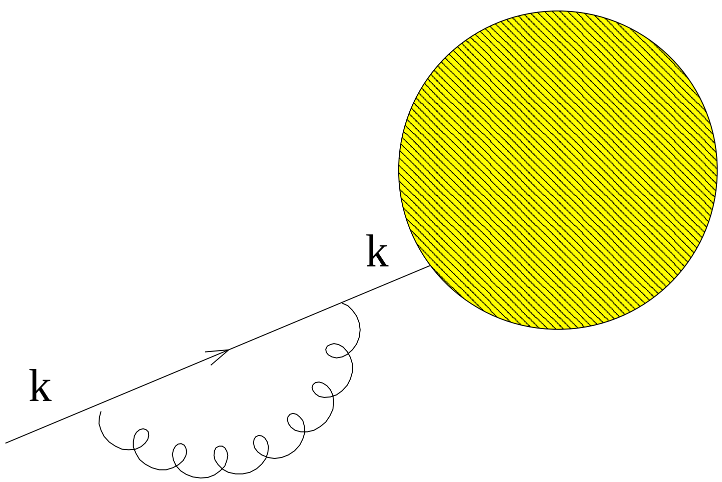
Consider now the case a process with an initial-state hadron: DIS

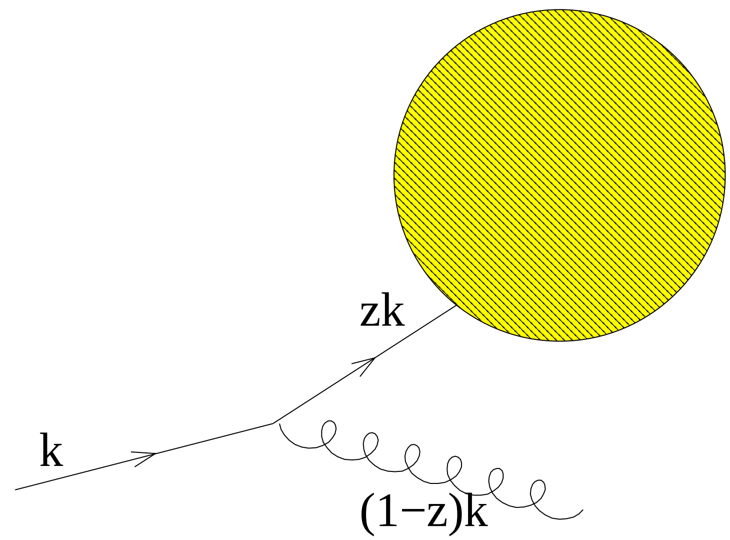
We know already the leading order: it's Feynman parton-model formula

$$d\sigma_{ep}(K) = \sum_q \int dx f_q(x) d\sigma_{eq}(xK)$$

with $d\sigma_{eq}$ the LO cross section for $eq \to eX$

Following what done before, we consider NLO corrections to $d\sigma_{eq}$





$$d\sigma_R + d\sigma_V = \frac{\alpha_S}{2\pi} \int dk_T^2 dz\, C_F \frac{1+z^2}{1-z} \frac{1}{k_T^2} \left(d\sigma^{(0)}(zk_a) - d\sigma^{(0)}(k_a) \right)$$

Finite for $z \to 1$ (soft), but *divergent* for $k_T \to 0$ (collinear)!

The real kinematic is not degenerate with the virtual one in the collinear limit. This does not happen in the case of final-state emissions

Tentative conclusion: the parton model *does not survive* radiative corrections

If so, pQCD can only be used for final-state hadrons

But there is a way out, which implies replacing the naive parton model by its QCD equivalent, the factorization theorem

Before going into that, a bit of notation

Plus distributions

Redefine the qq Altarelli-Parisi kernel as follows (a distribution)

$$P(z) = C_F \left(\frac{1+z^2}{1-z} \right)_+$$

This notation introduces the "plus prescription":

$$\int_0^1 dz h(z) (g(z))_+ = \int_0^1 dz (h(z) - h(1)) g(z)$$

The NLO corrections to the parton cross section can therefore be written in a much more compact form

$$d\sigma_R(k_a) + d\sigma_V(k_a) = \frac{\alpha_S}{2\pi} \int \frac{dk_T^2}{k_T^2} dz P(z) d\sigma^{(0)}(z k_a)$$

The $+$ reminds to subtract the $z = 1$ singularity $\Longleftarrow$ includes part of the virtual corrections

Recovering the parton model

Exclude the collinear divergence with a cutoff $\mu_0 \ll Q$. Inserting the partonic cross section into the parton model we get after the k_T integration

$$d\sigma^{(NLO)}(K) = \frac{\alpha_S}{2\pi} \log \frac{Q^2}{\mu_0^2} \int dydz f(y) P(z) d\sigma^{(0)}(yzK)$$

and with some algebra

$$d\sigma(K) \equiv d\sigma^{(0)}(K) + d\sigma^{(NLO)}(K) = \int dy \hat{f}(y, \mu^2, \mu_0^2) d\hat{\sigma}(yK, \mu^2, Q^2)$$

with $\mu_0 \ll \mu \sim Q$

$$\hat{f}(y, \mu^2, \mu_0^2) = f(y) + \frac{\alpha_S}{2\pi} \log \frac{\mu^2}{\mu_0^2} \int_y^1 \frac{dz}{z} P(z) f(y/z)$$

$$d\hat{\sigma}(K, \mu^2, Q^2) = d\sigma^{(0)}(K) + \frac{\alpha_S}{2\pi} \log \frac{Q^2}{\mu^2} \int_0^1 dz P(z) d\sigma^{(0)}(zK)$$

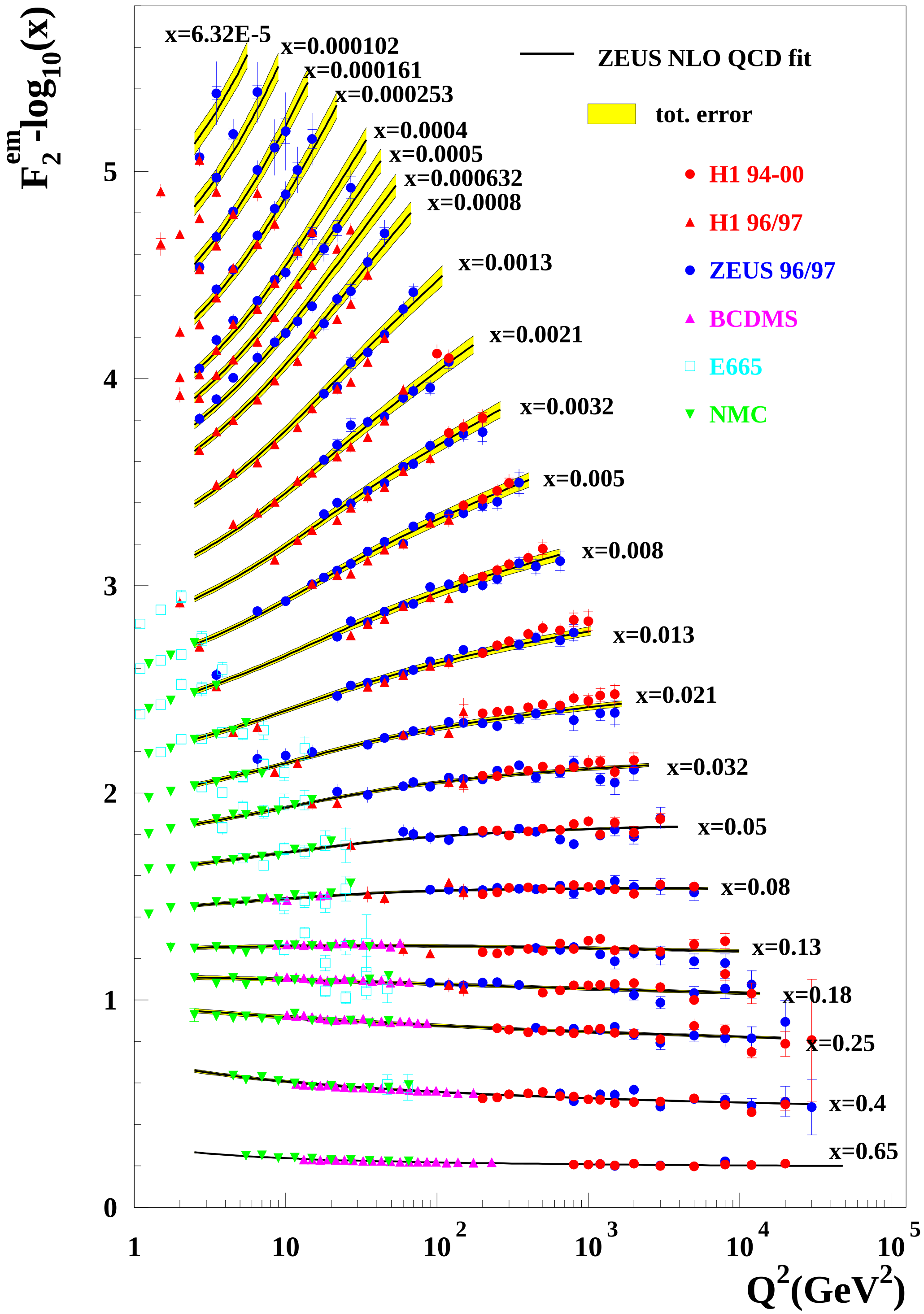
Note: it is $\hat{f}$ that is usually denoted by f

It is now manifest that the *divergence is independent of the process* (as for final-state emissions). Consequences

- PDFs acquire a dependence upon mass scales: scaling violations
- PDFs cannot be expanded in perturbation theory
- Parton cross sections do have a perturbative expansion

The key assumption: Nature will kill the $\log \mu_0$ divergence in the PDFs (smearing typical of long-distance phenomena). We cannot compute PDFs, but we can extract them from data

Parton model is formally recovered. An all-order proof of these QCD-improved formulae gives a *factorization theorem*



$$F_2^{\rm NC} = x \sum_f e^2(f) \left[q^{(f)} + \bar{q}^{(f)} \right] + \mathcal{O}(\alpha_S)$$

An excellent fit already at the NLO

If one derives the PDFs wrt the hard scale μ

$$\frac{\partial \hat{f}^{(H)}(y,\mu^2,\mu_0^2)}{\partial \log \mu^2} = \frac{\alpha_S}{2\pi}\int_y^1 \frac{dz}{z} P(z)\hat{f}(y/z,\mu^2,\mu_0^2) + \mathcal{O}(\alpha_S^2).$$

The cutoff dependence is entirely in $\hat{f}$ $\Longrightarrow$ sensible to assume that the r.h.s. is the first order of a well-behaved perturbative expansion

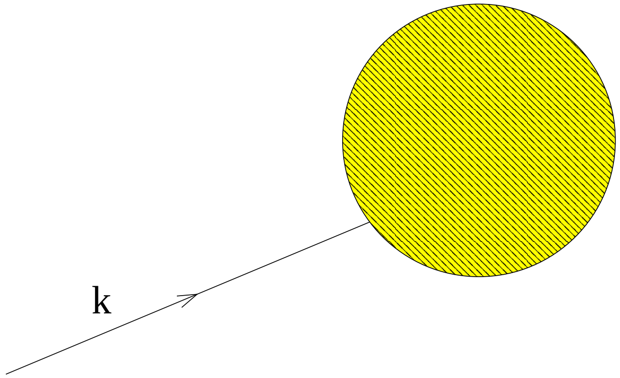
One therefore arrives at the Altarelli-Parisi equations (1977), being careful enough to include *all possible splitting* types

$$\begin{aligned} \frac{\partial \hat{f}_a}{\partial \log \mu^2} &= \sum_b P_{ab} \otimes \hat{f}_b \\ P_{ab} &= \alpha_S P_{ab}^{(0)} + \alpha_S^2 P_{ab}^{(1)} + \alpha_S^3 P_{ab}^{(2)} + \ldots \end{aligned}$$

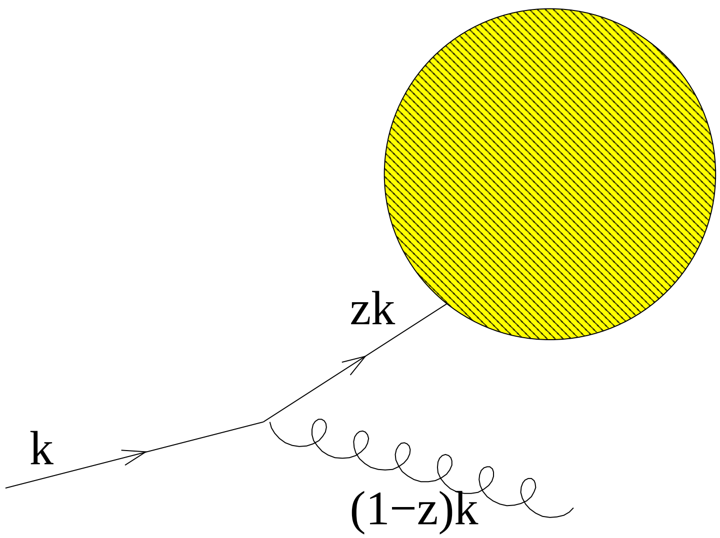
I introduced another frequently used notation

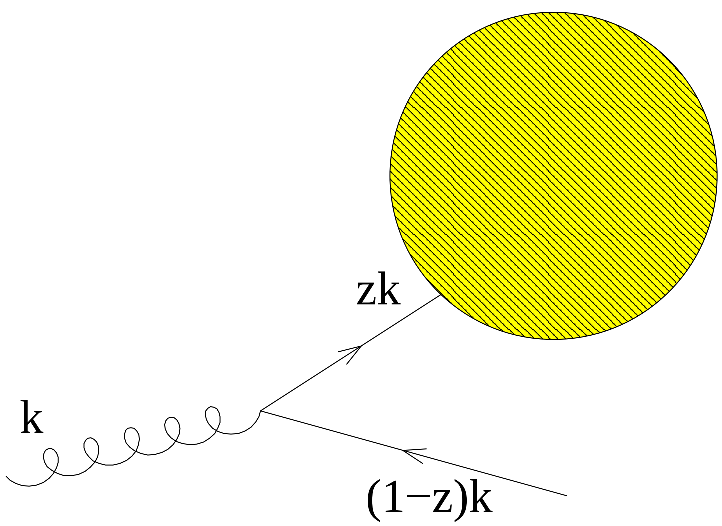
$$f = g \otimes h \quad \Longleftrightarrow \quad f(x) = \int_0^1 dy dz \delta(x - yz) g(y) h(z)$$

Note: the necessity of considering all splitting types is a consequence of perturbative corrections. The LO diagram



has the following real correction diagrams





History of AP kernels

- $P_{ab}^{(0)}$: Altarelli, Parisi (1977)
- $P_{ab}^{(1)}$: Curci, Furmanski, Petronzio (1980)
- $P_{ab}^{(2)}$: Moch, Vermaseren, Vogt (2004)

The calculation of $P_{ab}^{(2)}$ is the toughest ever performed in perturbative QCD, with 10^6 lines of dedicated algebraic code, and 20 man-year of work

- One loop $\Longrightarrow$ 18 Feynman diagrams
- Two loops $\Longrightarrow$ 350 Feynman diagrams
- Three loops $\Longrightarrow$ 9607 Feynman diagrams

We are on the right track for an exact determination of PDFs at the NNLO

Determination of PDFs

Key ingredients: factorization theorems and AP equations. Then (in Mellin space, where convolutions become ordinary products)

$$\sigma_{data} = f\sigma_{th} \implies f = \sigma_{data}/\sigma_{th}$$
$$\sigma_{data} = (f_1 f_2)\sigma_{th} \implies (f_1 f_2) = \sigma_{data}/\sigma_{th}$$

- Parametrize PDFs at a small scale $Q_0 = 1 - 4$ GeV

$$xf(x, Q_0) = Ax^{\delta}(1-x)^{\eta}(1 + \epsilon\sqrt{x} + \gamma x)$$

- Impose momentum conservation

$$\sum_a \int_0^1 dx x f_a(x, Q_0) = 1$$

- Evolve PDFs to relevant Q and compute σ_{th}
- Fit to data

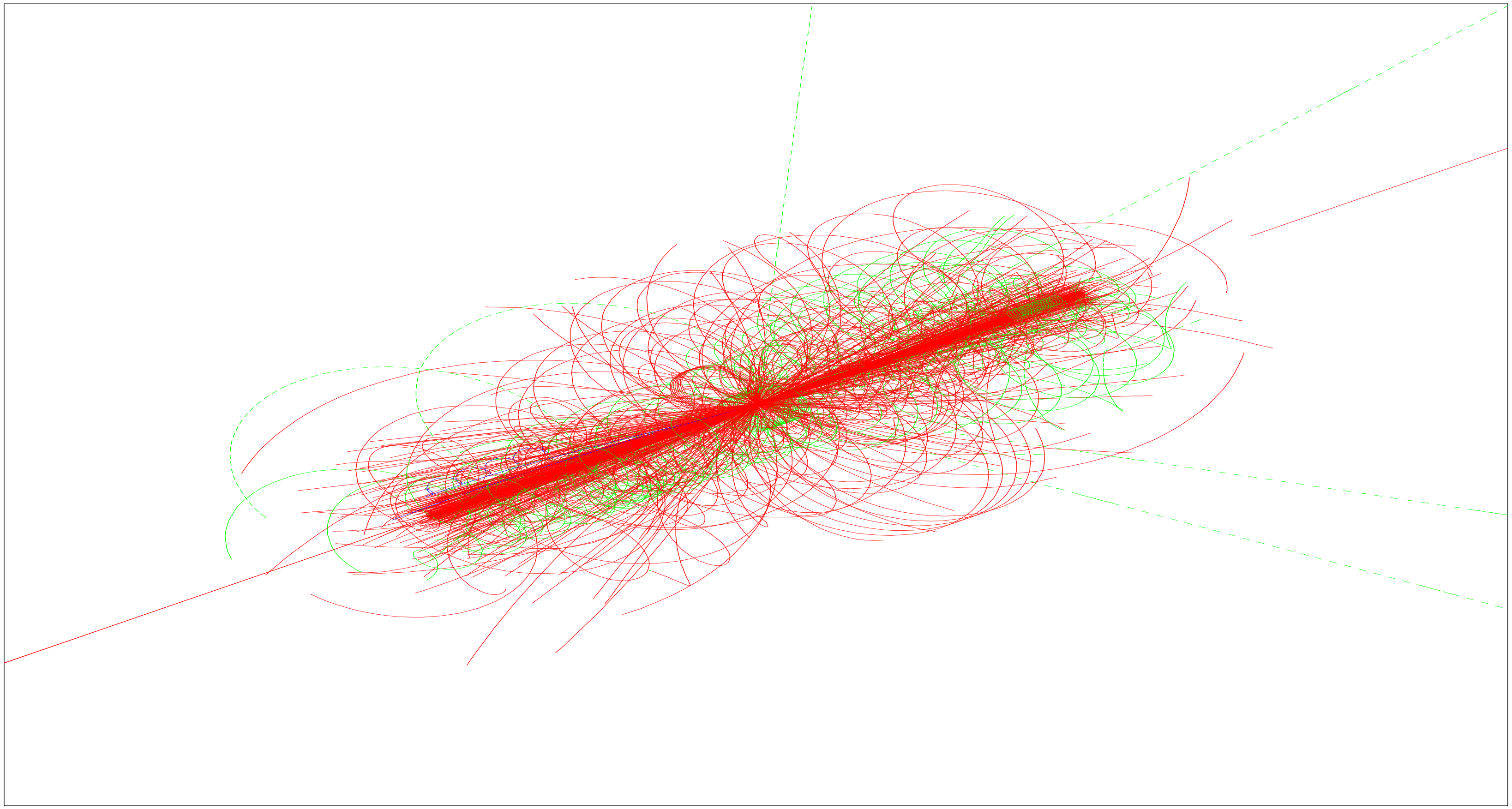
Summary on pQCD

- We are able to describe *some aspects* of a world of hadrons in terms of quarks and gluons
- The perturbative machinery works, if supplemented by non-perturbative inputs (PDFs)
- Intuitive ideas (parton model and factorization) survive in QCD, at the price of certain complications

This framework must be able to stand the challenge posed by data, and we can now say that it does it in an excellent way – the days of QCD *tests* are over; precision physics is possible

We can therefore confidently tackle the problem of predictings SM processes at the LHC. Which remains a very difficult problem...

A "typical" pp event



$H \to ZZ \to 4\mu$ as simulated by ATLAS

- Straight, dashed lines: μ's, i.e. the signal
- The rest: a big mess, due to the fact that hadrons are complicated objects

A complete description must account for two ingredients:

1) the hard process: all the high-p_T stuff, plus particles at small relative p_T or with small energies

2) the rest: this is generally low-p_T stuff, and includes

- the underlying event;
- the pile-up, ie other pp collisions

Truth be told, there's no unambiguous separation between 1) and 2), since to a certain extent it is always definition dependent

A complete description must account for two ingredients:

1) the hard process: all the high-p_T stuff, plus particles at small relative p_T or with small energies

2) the rest: this is generally low-p_T stuff, and includes

- the underlying event;
- the pile-up, ie other pp collisions

Two different approaches

♦ Event Generators: aim at giving a description as realistic as possible, including all the details of 1) and 2)

Examples: HERWIG, PYTHIA, ARIADNE, ...

♦ Cross Section Integrators: don't include 2), and are only able to give predictions for infrared-safe observables resulting from 1)

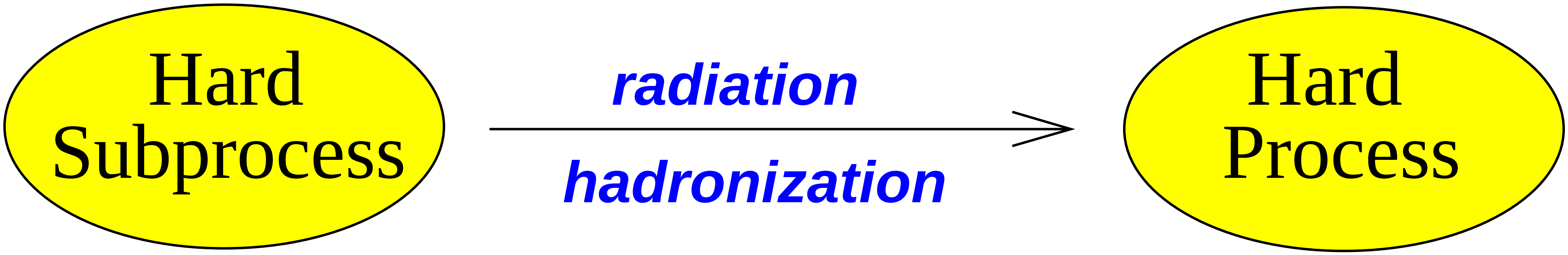
Examples: MCFM, ResBos, ...

A bit of terminology:

- Event Generators are frequently called Parton Shower Monte Carlos – not really correct, but not wrong either (we'll see why)

- Cross Section Integrators are called Monte Carlos (by theorists) – this is due to the fact that they use numerical monte carlo methods to carry out the necessary integrations

For both Event Generators and Cross Section Integrators, the simulation of the hard process proceeds schematically as follows



- Hard subprocess: *only* large-p_T particles, parton-level. Two partons pulled out of the incoming hadrons scatter and produce few (2–6) particles
- Radiation: adds more partons. Equivalent to considering *higher-order corrections* in perturbative QCD
- Hadronization: converts incoming partons into scattering hadrons, and outgoing partons into observed particles

Strategies

- For Hadronization
  1. Use factorization theorems $\longrightarrow$ Cross Section Integrators
  2. Use phenomenological models at mass scales where pQCD is not applicable $\longrightarrow$ Event Generators

- For Higher-order Corrections
  1. Compute exactly the result to a given order in α_S
  2. Estimate the dominant effects to all orders in α_S

  Cross Section Integrators may implement 1, 2, or a combination of the two. Event Generators always implement 2, possibly combined with 1

Summary so far

- It is convenient to separate high- from low-p_T phenomena
- High-p_T (ie hard) processes are *predicted*, low-p_T ones are *modeled* (and fitted to data)
- Cross Section Integrators will neglect the problem of low-p_T stuff if not associated with high-p_T particles
- Event Generators and CSIs both start from simulating a *hard subprocess*. They differ in the way radiation and hadronization are described

Cross Section Integrators

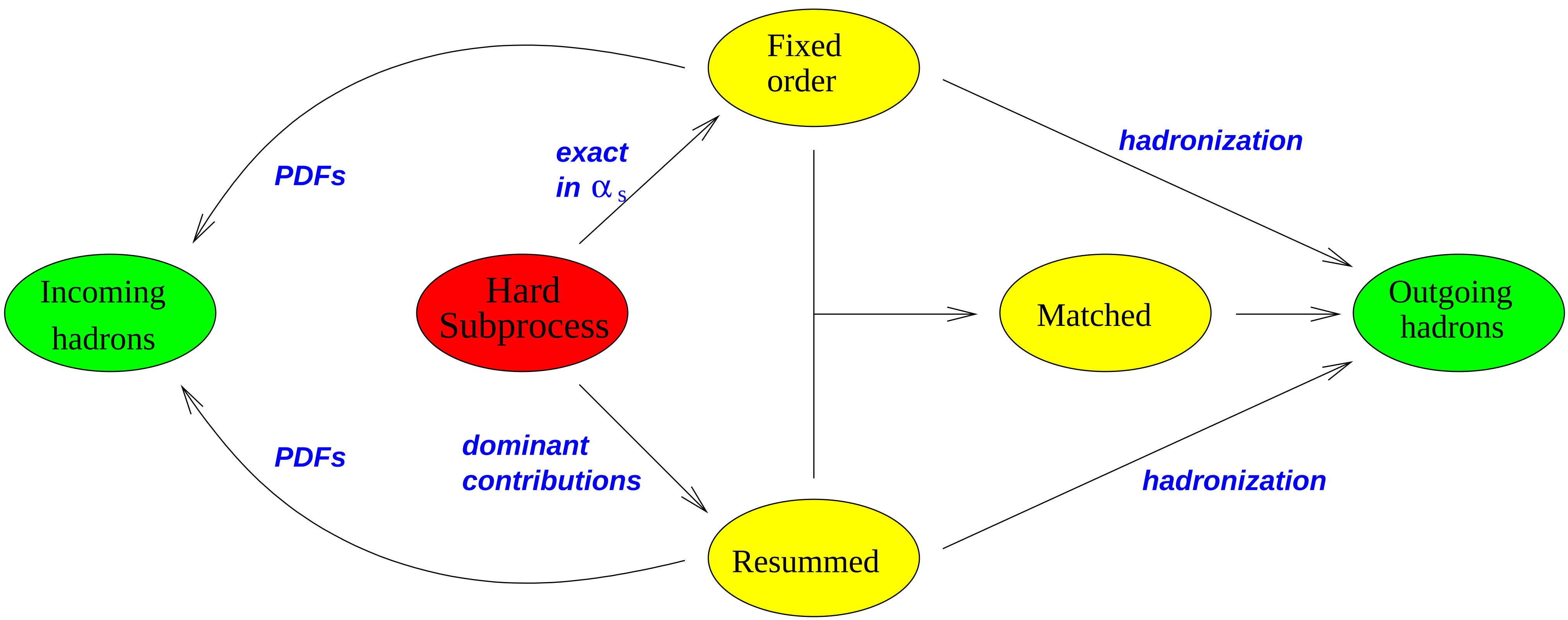
Keep in mind that

- CSIs do a good job in dealing with the hard process, computed using perturbative techniques
- CSI are basically *parton-level computations*
- Hadrons in the initial/final state are obtained by convoluting parton results with PDFS/fragmentation functions
- Unweighted unbiased events are in general not available beyond LO

CSIs can be broadly divided into two classes (which can be combined)

- Fixed-order (eg MCFM) $\longleftarrow$ exact to some α_S^k
- Resummation (eg ResBos) $\longleftarrow$ dominant effects to all orders in α_S

The making of the hard process with CSIs



Convolution with PDFs

The master formula is always the factorization theorem

$$d\sigma_{H_1H_2}(P_1,P_2) = \sum_{ij}\int dx_1 dx_2 f_i^{(H_1)}(x_1,\mu^2) f_j^{(H_2)}(x_2,\mu^2) \times d\hat{\sigma}_{ij}(x_1P_1,x_2P_2;\alpha_S(\mu^2),\mu^2)$$

In order to obtain a theoretical prediction, one computes $d\hat{\sigma}_{ij}$, then uses the formula above taking the PDFs from available repositories*

The classification of CSIs is equivalent to the classification of the short-distance parton cross sections $d\hat{\sigma}_{ij}$

Thus typically one deals with parton cross sections, understanding the convolution with the PDFs

* or to fit PDFs to data

Hadronization (fragmentation)

The idea: partons produced in the hard collision move fast away from each other. Each of them will eventually pick up (at large p_T) the missing colour and flavour from the vacuum to create an observable hadron

Example: b hadroproduction. The single-inclusive p_T spectrum of the b-flavoured hadron is:

$$\frac{d\hat{\sigma}_{ij\to H_b}}{dp_T(H_b)} = \int \frac{dz}{z} D^{b\to H_b}(z,\epsilon) \frac{d\hat{\sigma}_{ij\to b}}{dp_T(b)}, \qquad p_T(H_b) = z p_T(b)$$

- $d\hat{\sigma}_{ij\to H_b}$ is convoluted with the PDFs to get $H_1 H_2 \longrightarrow H_b$
- The *fragmentation function* $D^{Q\to H_Q}$ is analogous to the PDFs: it cannot be computed in pQCD, but is universal
- One tipically uses e^+e^- to fit the parameter(s) ϵ; the functional form in z must be guessed (Peterson, Kartvelishvili,...)

The hard subprocess may be seen as a zero-order approximation in the description of the hard process in CSIs

It is also useful to introduce in a simple manner:

- process kinematics
- candidate, weighted, and unweighted events

As an example, consider

$$H_1 H_2 \longrightarrow W + X$$

which gets contributions from the leading-order hard subprocesses

$$q\bar{q}' \to W \to e\nu$$

Kinematics of the hard subprocess

According to the factorization theorem

$$p_1 = x_1 P_1\,, \qquad p_2 = x_2 P_2\,, \qquad 0 < x_1, x_2 < 1$$

In the hadron c.m. frame (which is the lab frame)

$$p_1 + p_2 = E_{beam}(x_1 + x_2, 0, 0, x_1 - x_2)\,, \qquad E_{beam} = \sqrt{S}/2$$

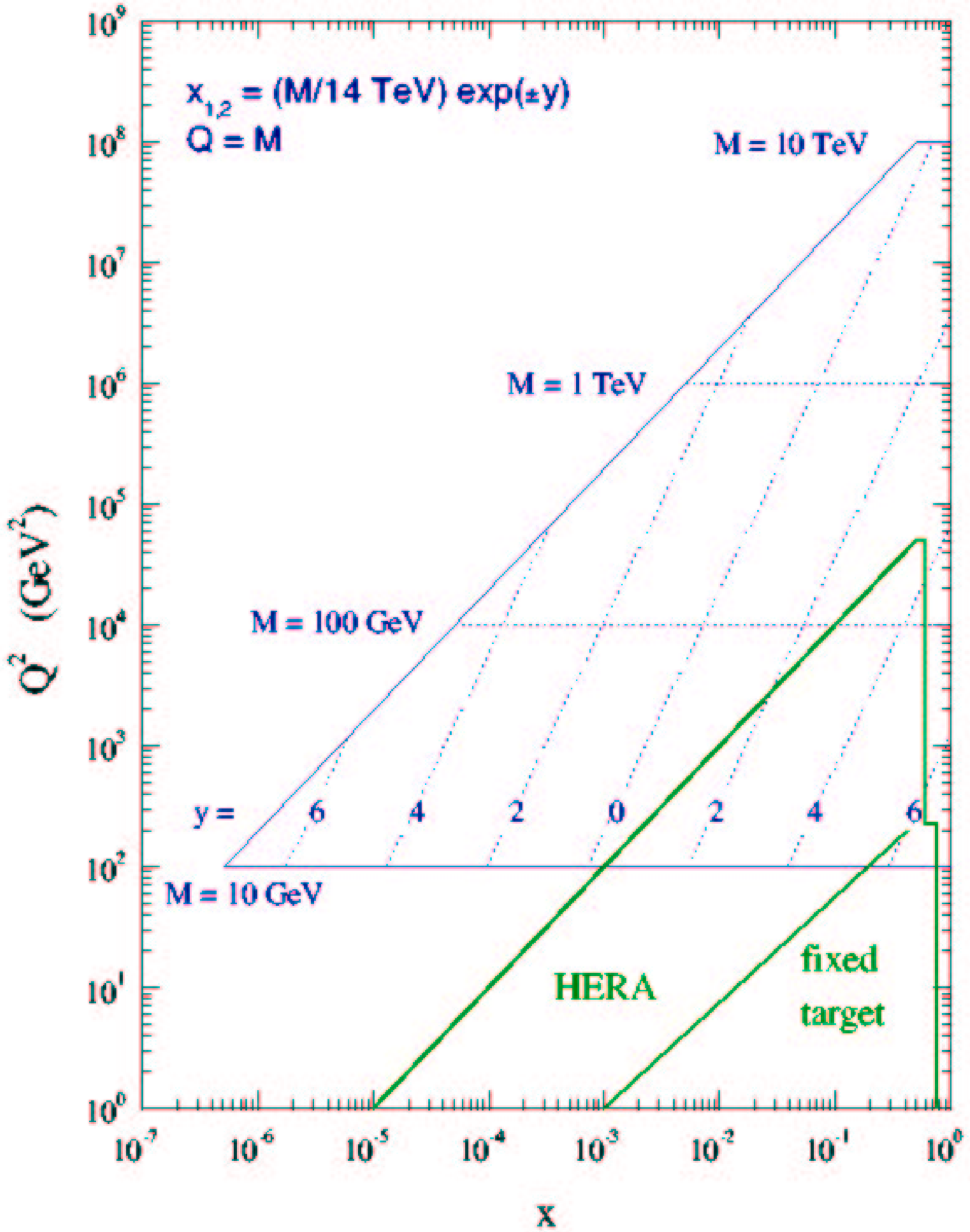
The parton c.m. frame is thus boosted wrt the lab frame by

$$y_{cm} = \frac{1}{2}\log\frac{(p_1+p_2)^0 + (p_1+p_2)^3}{(p_1+p_2)^0 - (p_1+p_2)^3} = \frac{1}{2}\log\frac{x_1}{x_2}$$

If the system produced in the collision has an invariant mass M and rapidity Y in the lab frame, $Y = y_{cm}$ and

$$(p_1+p_2)^2 = M^2 \quad \Longrightarrow \quad x_1 x_2 S = M^2$$

$$x_1 = \frac{M}{\sqrt{S}} e^{Y}\,, \qquad x_2 = \frac{M}{\sqrt{S}} e^{-Y}$$



x on the horizontal axis gives the kinematically allowed range for the ratio of parton energy over hadron energy, at fixed mass and rapidity of the system produced

The picture is surprisingly accurate, but not exact: QCD radiation will change it

In the context of factorization theorems, (x, Q^2) are the arguments of the PDFs

Hard subprocess events

A minimal and necessary information on the production process is given by the lowest-order cross section

$$d\sigma(q\bar{q}' \to W \to e\nu) = \frac{1}{2\hat{s}} \, |\mathcal{A}(q\bar{q}' \to W \to e\nu)|^2 \, \frac{d\cos\theta d\phi}{8(2\pi)^2} \, ,$$

Now sample the phase-space, ie pick up random values for the variables

$$(\cos\theta_i, \phi_i) \quad \Longleftrightarrow \quad (k_q + k_{\bar{q}'} \; \longrightarrow \; k_e + k_\nu)$$

This defines a *candidate event*

Candidate events do not correspond to anything observable

A *weighted event* is a candidate event, augmented by the candidate event's cross section

$$\begin{aligned} \mathcal{E}_i &= (w_i; \cos\theta_i, \phi_i)\,, \qquad w_i = V_\Phi \frac{d\sigma}{d\cos\theta d\phi}(\cos\theta_i, \phi_i) \\ V_\Phi &= \int d\cos\theta d\phi \end{aligned}$$

Weighted events are still not observable – their distribution is not the one observed in Nature, but that of the random number generation

However, by using their weights to fill histograms, one gets a faithful representation of measured spectra. The simplest example is the total rate

$$\langle w \rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} w_i = \int d\sigma$$

Weighted events can be used to predict observables

Unweighted events are the subset of candidate events which occur with the same frequency as that observed in Nature

Such subset can be obtained with the *hit-and-miss* technique (also known as Von Neumann method)

0) Scan the phase space and find an upper bound $w_{\max}$ for the cross section

$$V_\Phi \frac{d\sigma}{d\cos\theta d\phi} \leq w_{\max}$$

1) For each candidate event $(\cos\theta_i, \phi_i)$ generate a random number r

2) If

$$\frac{w_i}{w_{\max}} \geq r$$

keep the candidate event, else reject it

3) Iterate 1) and 2) N times

Through this procedure, unweighted events are distributed according to the hard cross section, and can therefore be given equal weight

$$w = \int d\sigma$$

It is also common to set $w = 1$

Note that in order to get M unweighted events, one has to generate N candidate events with

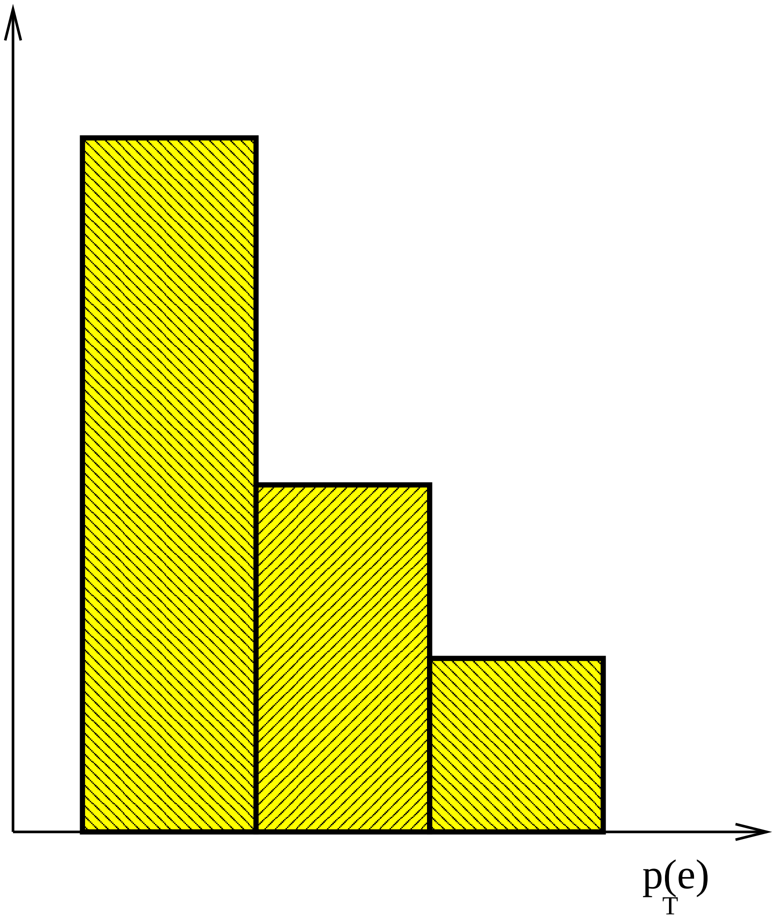
$$M \leq N, \qquad \varepsilon = \frac{M}{N} \leq 1$$

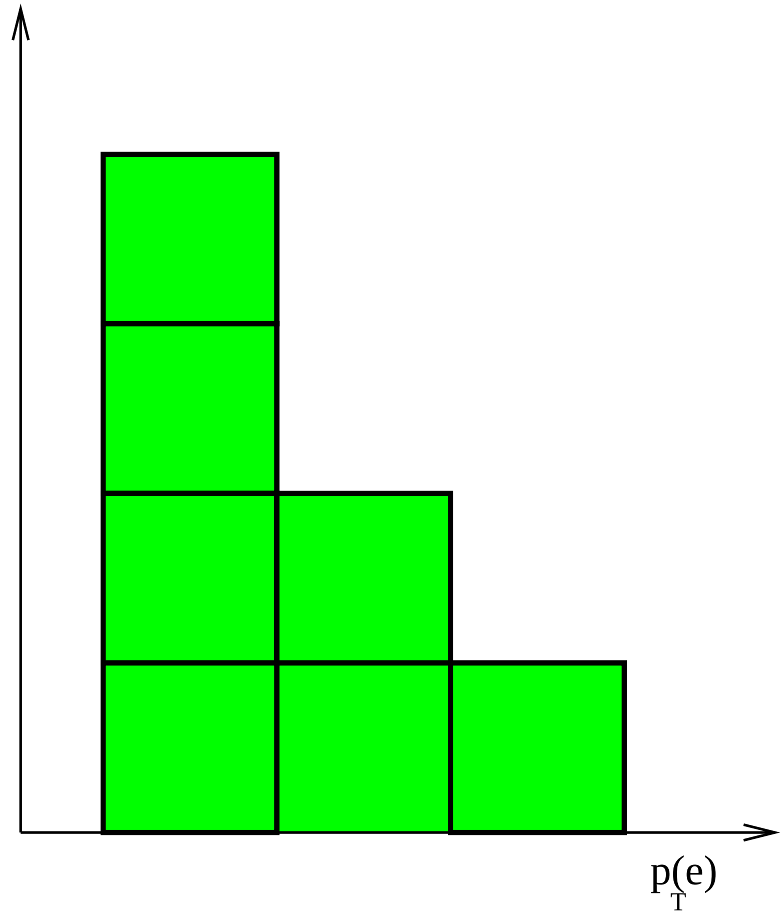
where ε is called *efficiency*

The efficiency is basically a function of $w_{\max}$ and of the distribution of the random numbers r. It is a figure of merit of the computation which achieves unweighted event generation

In summary

- Weighted events: equal probability*, unequal weights
- Unweighted events: unequal probability, equal weights





* If candidate events are generated flat

From hard subprocess to hard process

Through the inclusion of hadronization effects and higher-order corrections, the hard event is converted into the "physical" event, ie the best approximation of what happens in the detector according to the chosen method of computation

- More particles are present in the final state wrt the hard subprocess

  Still a small number, say less than 10 for CSIs

- It gives the W something to recoil against, and thus $p_T(W) > 0$

Higher-order corrections, however, pose a problem

- In the context of Cross Section Integrators, *unweighted events do not exist*. They can be defined only by introducing unphysical cutoffs, which bias observables

Unweighted physical events are only meaningful in Event Generators

Summary on CSIs

- Aim at giving an accurate description of hard processes
- Parton-level results (except for fragmented partons). Small final-state multiplicities (< 10)
- Unsuited for detector simulations. Best tools for "precision" tests, PDF extractions, α_S measurements
- May incorporate exact perturbative results up to α_S^k
  $\longrightarrow$ Fixed-order CSIs
- May incorporate approximate perturbative results to all α_S^n
  $\longrightarrow$ Resummed CSIs