

# Soft QCD parameter tuning using feed-forward neural networks

Michael Dessena

17th March 2022

## Abstract

—INSERIRE ABSTRACT—



# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Hadron-Hadron Scattering</b>	<b>7</b>
2.1	QCD factorization theorem . . . . .	7
2.2	Partonic cross section . . . . .	8
2.2.1	Higher order calculations . . . . .	8
2.3	Parton Showers . . . . .	11
2.3.1	Merging parton showers and matrix element calculations . . . . .	12
2.4	Parton distribution functions . . . . .	13
2.5	A real proton-proton collision . . . . .	13
<b>3</b>	<b>Multiple Parton Interactions</b>	<b>17</b>
3.1	Basic Concepts . . . . .	17
3.2	Pythia8 Monte Carlo events generator . . . . .	20
3.2.1	Parton Shower . . . . .	20
3.2.2	Multiple Parton Interactions in Pythia . . . . .	20
3.2.3	Momentum and flavour conservation . . . . .	21
3.2.4	Impact Parameter Dependence . . . . .	21
3.2.5	Parton rescattering . . . . .	22
3.2.6	Interleaving of Multiple Interaction and Parton Shower . . . . .	23
3.2.7	Beam Beam Remnants and primordial $k_T$ . . . . .	23
3.2.8	Color Reconnection and Hadronization . . . . .	25
3.3	PYTHIA summary . . . . .	26
<b>4</b>	<b>Observable to Study the Underlying Event</b>	<b>27</b>
4.1	Minimum Bias Measurements and Underlying Event topology . . . . .	27
<b>5</b>	<b>Tune procedure, CP5 Tune and MCNNTUNES</b>	<b>31</b>
	<b>Bibliography</b>	<b>33</b>



# Chapter 1

## Introduction

The *standard model of particle* is the theory that describe all the elementary particles, the electroweak [1] and the strong interaction



# Chapter 2

## Hadron-Hadron Scattering

In a high energy proton-proton collision we can have either soft or hard processes. Most of the time the hard processes are accompanied by soft interaction, occurring along the hadron interaction. The hard processes are well understood using perturbation theory, meanwhile the soft processes are not so well understood in fact these processes are described by non perturbative QCD.

### 2.1 QCD factorization theorem

The factorization theorem was introduced first by Drell and Yan [2].

The hadron-hadron scattering is describe in terms of partons extending the formalism used for deep inelastic scattering.

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a) f_{b/B}(x_b) \hat{\sigma}_{ab \rightarrow X} \quad (2.1)$$

Where X is a partonic/leptonic state and  $a(b)$  a quark or an antiquark in the hadron  $A(B)$ . This is valid in the "scaling" limit:

$$s \longrightarrow \infty \quad \frac{M_X}{s} = \text{finite} \quad (2.2)$$

The problem arise from the perturbative corrections from real and virtual gluon emission. This contribution give a logarithmic divergence (spoil the convergence of the perturbative expansion). These dependencies can be absorbed by the parton distribution function (DGLAP equations). This result in the violation of scaling:

$$f_{a/A}(x_a) \longrightarrow f_{a/A}(x_a, Q^2) \quad (2.3)$$

Now the parton distribution function depend on the momentum scale  $Q^2$ .

So, we can rewrite the factorization theorem in Eq. 2.1 as:

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) \hat{\sigma}_{ab \rightarrow X} \quad (2.4)$$

Now the *finite* corrections in the perturbative expansion are specific for each process (not universal). This leads in the Eq. 2.4 to the  $\alpha_s$  series:

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a, \mu_F^2) f_{b/B}(x_b, \mu_F^2) [\hat{\sigma}_0 + \alpha_s(\mu_R^2) \hat{\sigma}_1 + \dots] \quad (2.5)$$

In Eq. 2.5 two scales enter the formula:

- The *factorization scale*  $\mu_F$ : this scale separates long- and short- distance.



- The *renormalization scale*  $\mu_R$ : the scale at which is evaluated the strong coupling  $\alpha_s$ .

The higher-order corrections eliminate the cross section prediction dependencies on  $\mu_R$  and  $\mu_F$ . Typically, the scales are assumed to be equal. For example in the Drell-Yan process the standard choice is  $\mu_F = \mu_R = M$ , with  $M$  the lepton pair mass [3]. Other cases are the invariant masses of  $Z$ -boson and top quark or the jet transverse energy to study [3] the production cross sections for  $Z$ -bosons, top quarks and large E T jets.

The parton distribution functions used in the hard scattering are solution of the DGLAP (Dokshitzer–Gribov–Lipatov–Altarelli–Parisi) equation [4, 5, 6, 7]

$$\mu_F^2 \frac{\partial f_{i/H}(x, \mu_F^2)}{\partial \mu_F^2} = \sum_j \frac{\alpha_s(\mu_F^2)}{2\pi} \int_x^1 \frac{dz}{z} P_{i \rightarrow j}(z) f_{j/p}\left(\frac{x}{z}, \mu_F^2\right) \quad (2.6)$$

Where  $P_{i \rightarrow j}$  are the splitting functions: they are the probability to have a parton of type  $i$  that became, by the emission of a quark or a gluon, a parton  $j$ , carrying fraction  $z$  of the momentum of  $i$ .

The splitting functions have perturbative expansions:

$$P_{i \rightarrow j}(x, \alpha_s) = P_{i \rightarrow j}^{(0)}(x) + \frac{\alpha_s}{2\pi} P_{i \rightarrow j}^{(1)}(x) + \dots \quad (2.7)$$

This procedure has been used to calculate Standard Model cross section in  $p\bar{p}$  and  $pp$  scattering respectively at Tevatron and LHC energies as shown in Fig. 2.1.

The parton distribution functions dependence on  $Q^2$  can be derived theoretically via the DGLAP equations. While, the  $x$  dependence is given fitting the deep-inelastic and other hard-scattering processes experimental data. The experimental coverage in  $(x, Q^2)$ -plane is shown in Fig. 2.2 where is also underlined the relationship between  $(x, Q^2)$  and kinematic variables in Drell-Yan processes for a final state with invariant mass  $M$  and rapidity  $y$  is shown, further details in section 2 of [3]. Assuming that the factorization scale  $Q$  is equal to  $M$  (The reference center of mass energy is 13 TeV). So, for two incoming particles with four-momentum respectively  $p_1$  and  $p_2$  the relations with  $y$  and  $M$  are:

$$\begin{aligned} p_1^\mu &= \frac{\sqrt{s}}{2}(x_1, 0, 0, x_1) \\ p_2^\mu &= \frac{\sqrt{s}}{2}(x_2, 0, 0, -x_2) \end{aligned} \implies x_1 = \frac{M}{\sqrt{s}}e^y \quad x_2 = \frac{M}{\sqrt{s}}e^{-y} \quad (2.8)$$

where  $s = (p_1^\mu + p_2^\mu)^2$ .

## 2.2 Partonic cross section

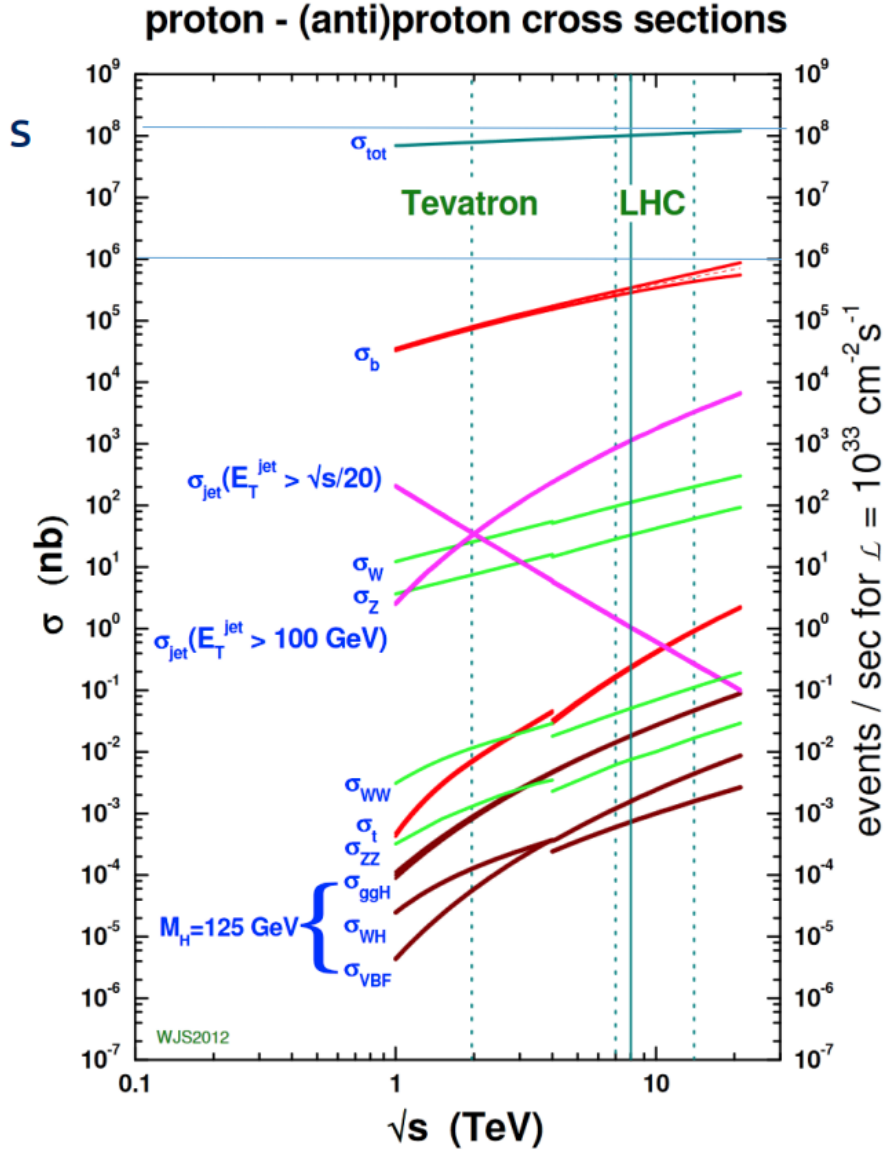
Partonic cross section is one of the fundamental ingredients in our recipe. This can be calculated in a perturbative series in  $\alpha_s$  from QCD first principles using quantum field theory.

The calculation at the leading order (LO) is performed evaluating all the possible tree-level Feynman diagrams for every process. Evaluating the squared matrix element and integrating over the available phase space (analytically or numerically).

Already here, we can encounter some divergence that have to be avoided imposing restriction on the phase space.

### 2.2.1 Higher order calculations

The LO calculation can describe broad feature of a particular process and provide a first estimation of its cross section but in many cases this is insufficient.



**Figure 2.1:** Next-to-leading order cross sections at Tevatron and LHC colliders energies (the splitting are at the transition between  $p\bar{p}$  and  $pp$  cross section). Figure from [8]

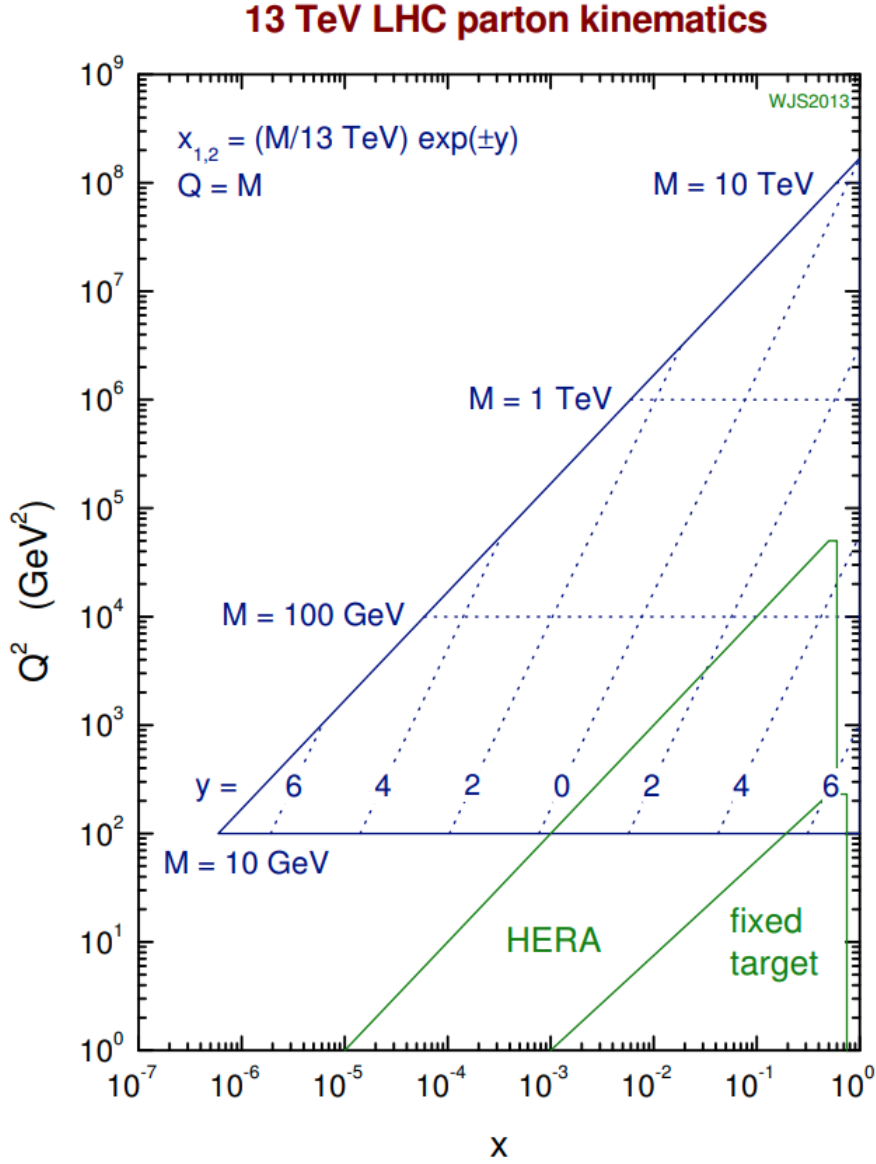
The main source of uncertainty derives from the LO dependence on the unphysical renormalization and factorization scales. Some process may contribute only when going beyond the first approximation, and some divergence can be resummed.

To the next-to-leading order (NLO) calculation participate all the Feynman diagrams that take an extra  $\alpha_s$ . This contribution can arise in two different way:

- Virtual: internal lines (loops);
- Real: external lines (real particles).

Virtual corrections contains infrared divergences, arising from the integral on the loop circulating momentum, that cancel against infrared singularities given by collinear emissions or soft emissions [9, 10, 11].

A common strategy for the renormalization is dimensional regularization: consists into perform the calculation in a  $D = 4 - 2\epsilon$ -dimensional space ( $\epsilon < 0$ ), in that way the singularities appear as single and double poles in  $\epsilon$ . Then, the limit  $\epsilon \rightarrow 0$  is taken after the divergences have cancelled.



**Figure 2.2:** Graphical representation of the parton  $(x, Q^2)$  variables coverage of different experiments. For LHC these  $x$  and  $Q^2$  are related to the kinematic variables  $y$  and  $M$ . Figure from [8]

This NLO calculation with regularization allows to extend the treatment to zero transverse momentum.

The importance of higher order calculations can be seen with the following example. In a  $Z$  boson production:

- 1) **LO**: the  $Z$  is produced without transverse momentum ( $p_T$ ), anything can recoil against the  $Z$  for momentum conservation.
- 2) **NLO**: the  $Z$  acquire a finite  $p_T$ , in this case the  $Z$  boson  $p_T$  is balanced by a single parton/gluon.
- 3) **NNLO**: the  $Z$   $p_T$  can be balanced by two jets.

Another important benefit of performing a NLO calculation is the the reduction of the dependence on the unphysical renormalization ( $\mu_R$ ) and factorization ( $\mu_F$ ) scales. It is proven that higher order calculations is that observables calculated to order  $\alpha_s^n$  are dependent on the

unphysical scales only at order higher than  $\alpha_s^{n+1}$  [3]. The range of predictions corresponding to different scale choices is usually attributed to *theoretical uncertainties*.

## 2.3 Parton Showers

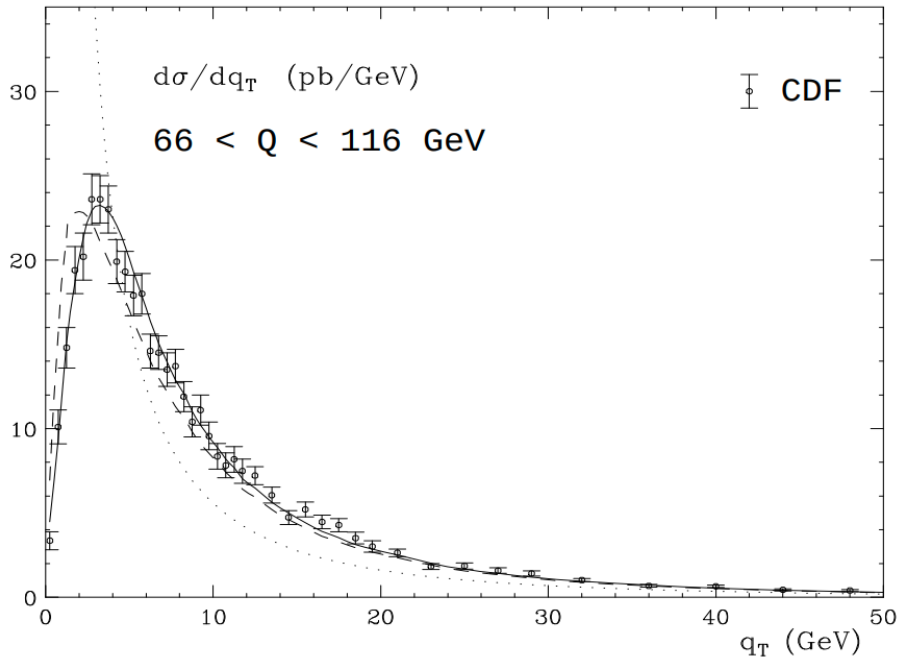
A different approach, instead than calculating order by order in the perturbative expansion, is to use an *all-order* approach in order to describe the phenomena observed at high-energy colliders.

different all-order approaches exist e.g. resummation and parton showers. Resummation is based on the observation that in many quantities the smallness of the expansion coefficients  $\alpha_s$  is violated by large logarithmic enhancements. This take the dominant contribution from each order and "resum" them by means of an evolution equation. The main problem in QCD is related to the fact that lot of quantities have correction of the form  $\alpha_s^n \log^k(Q_i/Q_j)$  where  $Q_i$  and  $Q_f$  are two different energies scales, for example:

- Renormalization and factorization scales logs:  $\alpha_s^n \log^n(Q^2/\mu_f)$

Various methods to perform this resummation exist.

An example is the  $Z$  production  $p_T$  spectrum shown in Fig. 2.3. A comparison between experimental CDF data and theoretical predictions is shown: in the low  $p_T$  region the *all-order* approach regularize the divergence of the fixed order calculation and describe the data better then the last one.



**Figure 2.3:** CDF data on  $Z$  production cross section at Tevatron collider, CDF experiment, the predictions from fixed order calculation (dotted) with resummation (dashed), and with the inclusion of power corrections (solid) are compared. Figure take from [12]

An other *all-order* approach is parton showers, this is implemented in different programs PYTHIA [13], HERWIG [14] and SHERPA [15]. This start from few parton arising from hard interaction and than these are related to partons to a lower energy scale close to  $\Lambda_{QCD}$  using the DGLAP evolution equation formalism, the solution at this equation can be write using a

Sudakov form factor arising from the probability of no gluon emission in the evolution from higher scale to lower scale.

In the parton showering process, additionally to the kinematic variable (momentum fraction  $z$  and an azimuthal angle  $\phi$ ) and flavours of the partons, an evolution variable  $t$  is generated. PYTHIA8 use as evolution variable the squared of the relative transverse momentum of the two partons in the splitting ( $p_T^2$ ). Different choices are made in HERWIG and SHERPA.

So as mentioned before, the shower evolution is based on the standard (LO) DGLAP splitting kernels  $P(z)$

$$P_{q \rightarrow qg}(z) = C_F \frac{1+z^2}{1-z} \quad (2.9)$$

$$P_{g \rightarrow gg}(z) = C_A \frac{(1-z(1-z))^2}{z(1-z)} \quad (2.10)$$

$$P_{q \rightarrow q\bar{q}}(z) = T_R(z^2 + (1-z)^2) \quad (2.11)$$

where  $C_F = \frac{4}{3}$ ,  $C_A = N_C = 3$  and  $T_R = \frac{1}{2}$  multiplied by  $N_f$  if summing over all contributing quark flavours.

Both Initial State Radiation (ISR) and Final State Radiation (FSR) algorithms are based on these splitting kernels.

The probabilities of emitting radiation as one moves in the decreasing evolution variable sequence are:

$$FSR : \quad \frac{d\mathcal{P}_{FSR}}{dp_T^2} = \frac{1}{p_T^2} \int \frac{dz}{z} \frac{\alpha_s}{2\pi} P(z) \quad (2.12)$$

$$ISR : \quad \frac{d\mathcal{P}_{ISR}}{dp_T^2} = \frac{1}{p_T^2} \int \frac{dz}{z} \frac{\alpha_s}{2\pi} P(z) \frac{f'(x/z, p_T^2)}{f(x, p_T^2)} \quad (2.13)$$

We can write our Sudakov form factor as:

$$\Delta(p_T^2) = \exp \left( - \int_{p_{T0}}^{p_T'} \frac{d\mathcal{P}_{PS}}{dp_T^2} dp_T \right) \quad \text{with} \quad PS = ISR, FSR \quad (2.14)$$

So, the Sudakov form factor give the probability of a parton to evolve from an harder scale to a softer scale without emitting a parton harder than some resolution scale.

The introduction of the Sudakov form factor resums all the effects from the soft and collinear gluon emission. For more details and some plots of different Sudakov form factor values see section 3.5 of [3].

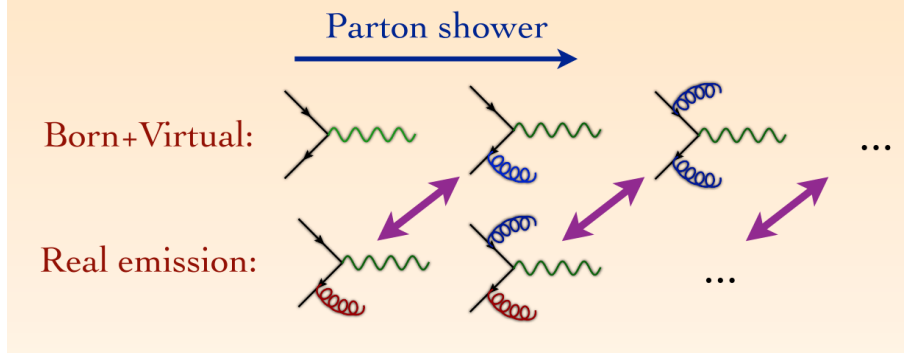
### 2.3.1 Merging parton showers and matrix element calculations

So, regions dominated by soft and collinear gluon emission are described very well by parton showers approach, on the other hand regions where partons are energetic and widely separated are well described by matrix element calculations.

The best approach would be to combine the two different description, This require an universal formalism for parton showers and matrix element calculations. This was created in 2001 and it is call "Les Houches Accord" [16]. In order to combine the two approach some care must be taken: there is the risk of double counting. There are different technique that prevent this risk for example CKKW [17].

A more best way is to combine NLO matrix element calculation with parton showers this is done by Frixione, Nason, Webber in the MC@NLO framework [18, 19, 20]. In this scenario the double counting is given by the fact that at NLO one emission is made real and the progress of

the parton shower give a double counting between real and virtual emission as shown in Fig. 2.4



**Figure 2.4:** See <https://indico.cern.ch/event/221788/attachments/361954/503905/frederix.pdf>

## 2.4 Parton distribution functions

The information on the quark distribution inside an hadron  $f_{q/p}(x, Q^2)$  arise from lepton-hadron DIS experiments and from lepton-pair production in hadron-hadron collisions (Drell-Yan processes) and jet measurements to study gluon distribution  $f_{g/p}(x, Q^2)$ . All these quantities are the experimental input in order to evaluate the PDF inside the hadron while the  $Q$  evolution can be described by DGLAP equation. The evolution of the PDF can be run either with a NLO calculation or with a NNLO calculation.

The kinematic region covered by experiments was shown in Fig. 2.2. At very low  $x$  and  $Q^2$  the DGLAP evolution is believed to be no longer applicable and a BFKL [ADD citations from [3] page 44] description must be used, but there are no evidence of this so the DGLAP approach is use as default in all the PDF analysis.

A lot of process are available for the PDFs evaluation and a lot of PDF set have been generated for example Fig. 2.5 shown the NNPDF3.1 set [21] at NNLO for a virtuality  $Q^2 = 10 \text{ GeV}^2$  (left) and  $Q^2 = 10^4 \text{ GeV}^2$  (right).

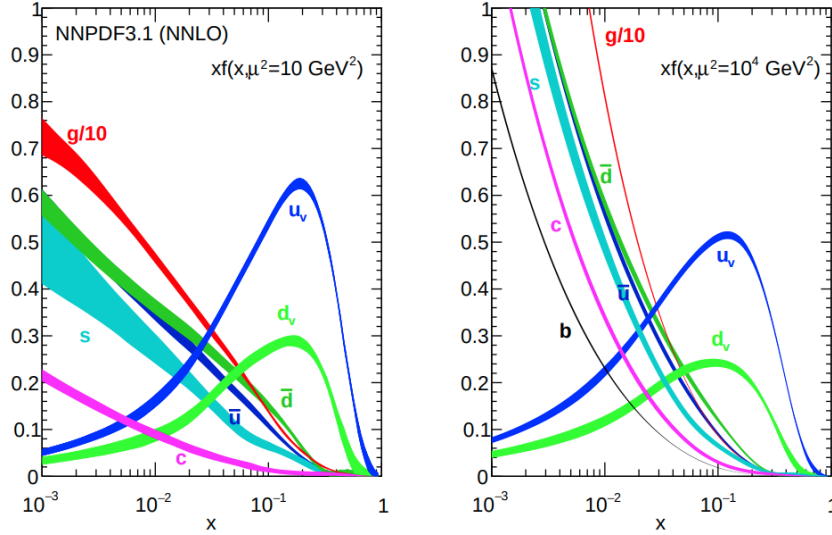
Note that the gluon contribution have been scaled of a factor 10 in fact in the low  $x$  region,  $x < 0.01$ , the gluon contribution is the dominating one. While, at high  $x$  value the valence quarks dominate the PDF. In Fig. 2.5 we can also see that with increasing virtuality ( $Q^2$ ) at low  $x$  the density of the sea quarks increase this is related to the fact that our hadrons are probed at higher energy and the probe resolution is proportional to the energy.

$$\text{Resolution} \sim \frac{\hbar}{Q^2} \quad (2.15)$$

So, when probed at higher energy the hadron appears denser than when are probed at lower energy. This, as will be discussed in the next chapter, is related to the higher number of interaction between partons in a single hadron-hadron collision. The phenomenon of having more than one interaction between partons in a single hadron-hadron collision is called multiple parton interaction and we want to introduce this concept in the next chapter.

## 2.5 A real proton-proton collision

We have understood that the complexity in the description of a proton-proton collision arise from composite nature of the protons. In this chapter we discussed the importance of the QCD factorization theorem that help us in the calculation of the hadronic cross section with the



**Figure 2.5:** The NNPDF3.1 NNLO PDFs, evaluated at  $Q^2 = 10 \text{ GeV}^2$  (left) and  $Q^2 = 10^4 \text{ GeV}^2$  (right). At low  $x$  the contribution from the gluon is the dominating one while at higher  $x$  the dominant contribution is from the valence quarks. The PDF  $Q$  evolution shows that when our proton is probed to higher  $Q^2$  the resolution increases (higher  $Q$  corresponds to smaller distance resolution) and so we have a bigger contribution from the sea quarks at low  $x$  values.

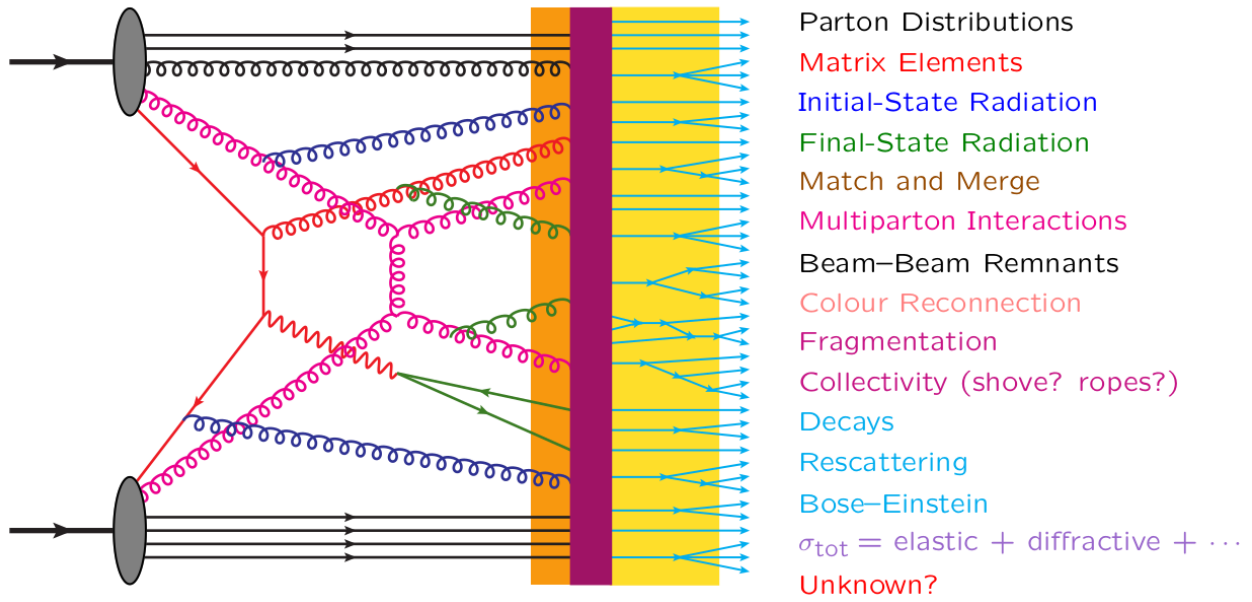
convolution between the partonic cross section and the PDF. We have discussed the importance of the parton shower algorithm where a set of partons are evolved in a more complex final state by emissions in the initial and final states.

All these processes are important in the description of a real proton-proton collision but also the partons that are left unscattered are non-color singlet and contribute to the complex final state observed in the experiment, and additionally, as mentioned before, nothing prevents additional parton scatterings from taking place and growing more and more the complexity of the partons final state.

Then, another problem is related to the not-well-understood hadronization process. Hadronization is not known from first principles and different models are implemented in different programs: the *cluster fragmentation model* implemented in HERWIG and the *string fragmentation model* in PYTHIA simulated this process of hadronization where the set of final-state partons are transformed into a set of hadrons. All these processes are shown schematically in Fig. 2.6.

Next chapter describes the PYTHIA Monte Carlo generator in more details. PYTHIA introduces different free parameters that need to be tuned with experimental data from Tevatron and LHC. The tune methods are described in Chapter 5





**Figure 2.6:** A schematic representation for a  $pp$  collision. Reading the image from left to right one can have an idea on the evolution of the system. The two incoming hadrons enter the scattering from the left side, the red line indicate the main hard scattering and the fuchsia one the second parton scattering (MPI) each interaction is associated with initial (blue) and final (green) state radiation, the unscattered partons (black lines) re-enter the color reconnection and hadronization processes. Then the new formed hadrons (lightblue) can undergo to different decays.





# Chapter 3

## Multiple Parton Interactions

The fact that the hadrons are viewed as "bunches" of partons it is likely that in an hadron-hadron different couples of partons can undergo to a scattering. This phenomenon is known as Multiple Parton Interactions (MPI) and is related to the composite nature of the incoming hadrons.

So, at some level these MPI have to exist, and are important in the description of the event: They can change the color topology of the colliding system as a whole,

In this scenario it is important to have a good understanding of the phenomenon. The aim of this section is to describe the basic concepts that are used to simulate MPI, for example in Pythia Monte Carlo event generator.

### 3.1 Basic Concepts

The main hypothesis of the multiple interaction models is that the QCD factorization is true not only for the hard process but also for the other scatters.

So, we can write:

$$\frac{d\sigma_{\text{int}}}{dp_{\perp}} = \sum_{i,j,k,l} \int dx_1 \int dx_2 \int d\hat{t} f_i(x_1, Q^2) f_j(x_2, Q^2) \frac{d\hat{\sigma}_{ij \rightarrow kl}}{d\hat{t}} \delta\left(p_{\perp}^2 - \frac{\hat{t}\hat{u}}{\hat{s}}\right) \quad (3.1)$$

This is the differential cross section for QCD hard  $2 \rightarrow 2$  processes, these processes are the one reported in Table 3.1.

In Eq. 3.1 we use the Mandelstam variables associated to the partonic system:

$$\hat{s} = (p_1 + p_2)^2 = (p_3 + p_4)^2 = x_1 x_2 s \quad (3.2)$$

$$\hat{t} = (p_1 - p_3)^2 = (p_2 - p_4)^2 \quad (3.3)$$

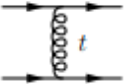

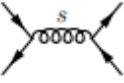
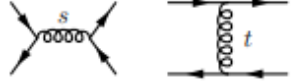
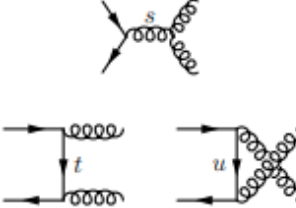
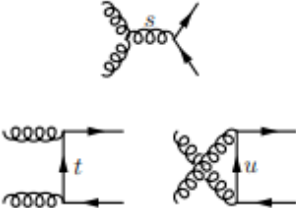
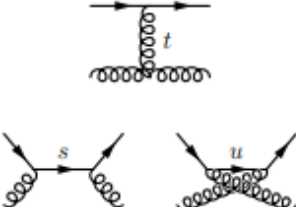
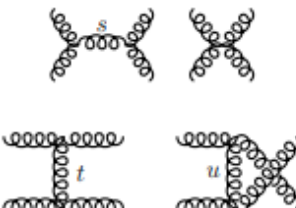
$$\hat{u} = (p_1 - p_4)^2 = (p_2 - p_3)^2 \quad (3.4)$$

where  $p_1, p_2$  are the four-momenta of the incoming partons and  $p_3, p_4$  the four-momenta of the outgoing partons. Note that in Eq. 3.1 the jet cross section is twice as large  $\sigma_{\text{jet}} = 2\sigma_{\text{int}}$ , because at first approximation each interaction gives rise to two jets.

We assume also that the "hardness" of processes is defined by the  $p_{\perp}$  scale of the interaction ( $Q^2 = p_{\perp}^2$ ).

As you can see from the formulae in Table 3.1 at small scattering angles: for  $t \rightarrow 0$ , the t-channel gluon exchange processes  $qq' \rightarrow qq'$ ,  $qg \rightarrow qg$  and  $gg \rightarrow gg$  dominate the full matrix element. For scattering that are soft relative to  $\hat{s}$ ,  $|\hat{t}| \ll \hat{s}$ , we can approximate  $|\hat{t}|$  as:

$$p_T^2 = \frac{\hat{t}\hat{u}}{\hat{s}} = \frac{\hat{t}(-\hat{s} - \hat{t})}{\hat{s}} \approx |\hat{t}| \quad (3.5)$$

Process	Amplitudes	$\overline{\Sigma}  \mathcal{M} ^2 / (4\pi\alpha_s)^2$
$q q' \rightarrow q q'$		$\frac{4}{9} \frac{s^2 + u^2}{t^2}$
$q q \rightarrow q q$		$\frac{4}{9} \frac{s^2 + u^2}{t^2} + \frac{4}{9} \frac{s^2 + t^2}{u^2} - \frac{8}{27} \frac{s^2}{tu}$
$q \bar{q} \rightarrow q' \bar{q}'$		$\frac{4}{9} \frac{t^2 + u^2}{s^2}$
$q \bar{q} \rightarrow q \bar{q}$		$\frac{4}{9} \frac{s^2 + u^2}{t^2} + \frac{4}{9} \frac{t^2 + u^2}{s^2} - \frac{8}{27} \frac{u^2}{st}$
$q \bar{q} \rightarrow g g$		$\frac{32}{27} \frac{t^2 + u^2}{tu} - \frac{8}{3} \frac{t^2 + u^2}{s^2}$
$g g \rightarrow q \bar{q}$		$\frac{1}{6} \frac{t^2 + u^2}{tu} - \frac{3}{8} \frac{t^2 + u^2}{s^2}$
$q g \rightarrow q g$		$-\frac{4}{9} \frac{s^2 + u^2}{su} + \frac{s^2 + u^2}{t^2}$
$g g \rightarrow g g$		$\frac{9}{2} \left( 3 - \frac{tu}{s^2} - \frac{su}{t^2} - \frac{st}{u^2} \right)$

**Table 3.1:** Parton-Parton differential cross sections ( $2 \rightarrow 2$  QCD process), can be calculated in pQCD evaluating the matrix element for each process involving quark, antiquark and gluon. Table from [22]

In this limit the only difference between quark and gluon are the color factors:

$$\hat{\sigma}_{gg} : \hat{\sigma}_{qg} : \hat{\sigma}_{qq} = \frac{9}{4} : 1 : \frac{4}{9} \quad (3.6)$$

So the Eq. 3.1 can be rewritten as:

$$\frac{d\sigma_{int}}{dp_T^2} \approx \int \int \frac{dx_1}{x_1} \frac{dx_2}{x_2} F(x_1, p_T^2) F(x_2, p_T^2) \frac{d\hat{\sigma}_{2 \rightarrow 2}}{dp_T^2} \quad (3.7)$$

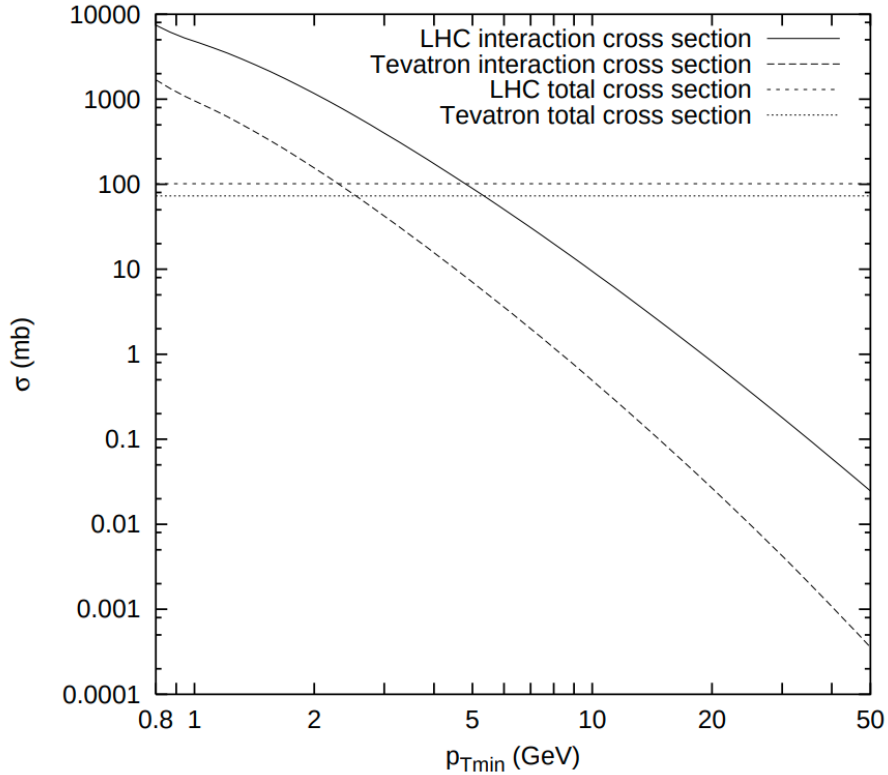
whit:

$$\frac{d\hat{\sigma}_{2 \rightarrow 2}}{dp_T^2} = \frac{8\pi\alpha_s^2(p_T^2)}{9p_T^4} \quad (3.8)$$

$$F(x, Q^2) = \sum_q (x q(x, Q^2) + x \bar{q}(x, Q^2)) + \frac{9}{4} x g(x, Q^2) \quad (3.9)$$

So we can integrate the Eq. 3.7:

$$\sigma_{int}(p_{Tmin}) = \int_{p_{Tmin}^2}^{(\sqrt{s}/2)^2} \frac{d\hat{\sigma}_{2 \rightarrow 2}}{dp_T^2} dp_T^2 \propto \frac{1}{p_{Tmin}^2} \xrightarrow{p_{Tmin} \rightarrow 0} \infty \quad (3.10)$$



**Figure 3.1:** This figure shows the interaction cross section ( $\sigma_{int}$ ) at Tevatron ( $p\bar{p}$ ,  $\sqrt{s} = 1.8$  TeV) and at LHC ( $pp$ ,  $\sqrt{s} = 14$  TeV). The flat lines are the corresponding values for the total cross section.

The total cross section is divergent in the limit  $p_T \rightarrow 0$ , this divergence is shown in Fig. 3.1. Due to this divergence the total interaction cross section at some  $p_T$  scale can exceed the total proton-proton cross section.

- To understand this paradox should be noted that the interaction cross section described in Eq. 3.10 is related to the interaction probability between two partons and counts the number of interactions, while the total proton-proton cross section  $\sigma_{pp}$  counts the number of events. For example, an event (a proton-proton collision) in which two partons interact counts once in the total cross section and twice in the interaction cross section.

So the ratio between this two quantities is perfectly allowed to be larger than unity:

$$\frac{\sigma_{int}(p_{T\ min})}{\sigma_{tot}} = \langle n \rangle(p_{T\ min}) \quad (3.11)$$

Secondly, we have to consider the screening effect: in fact the incoming hadrons are color singlet objects. Therefore, when the  $p_T$  of an exchanged gluon is small, and so the associated wavelength large, this gluon can no longer resolve the color charges and the effective coupling is decreased.

This cutoff is associated with color screening distance i.e. the average size of the region within the compensation of color charge occurs. This cutoff is then introduced in the factor:

$$\frac{\alpha_s^2(p_{T0}^2 + p_T^2)}{\alpha_s^2(p_T^2)} \frac{p_T^4}{(p_{T0}^2 + p_T^2)^2} \quad (3.12)$$

This factor contains the phenomenological regularization of the divergence, with the factor  $p_{T0}$  that have to be tuned to data.

Now the interaction cross section i

This parameter  $p_{T0}$  do not have to be energy-independent. But the energy is related with the sensibility of our probe, higher energy is related with the capacity of probing PDF to lower  $x$  values (as discussed previously, see Fig. 2.5), this low  $x$  region the number of partons rapidly increases. So, the partons are closer packed in this regions and as a consequence the color screening distance  $d$  decrease.

The number of partons is related to  $x$  with a power law so it is likely to have a dependence of the same form for  $p_{T0}$  respect to the CM energy.

$$p_{T0}(\sqrt{s}) = p_{T0}^{ref} \left( \frac{\sqrt{s}}{E_{CM}^{ref}} \right)^{E_{CM}^{pow}} \quad (3.13)$$

## 3.2 Pythia8 Monte Carlo events generator

Pythia8 is a standard tool for the simulation of events in high energy collisions. Pythia contains the evolution from a few-body system to a complex multiparticle final state.

### 3.2.1 Parton Shower

In Pythia8 all the contributions from Initial State Radiation (ISR), Final State Radiation (FSR) and Multi Parton Interactions (MPI) are interleaved into a single common sequence of decreasing  $p_T$ .

The parton shower has been described in the previously section. The solution to the DGLAP equation is given putting Eq. 2.12 and Eq. 2.13 in Eq. 2.14 for the ISR and the FSR by a Sudakov form factor that is related to the no emission probability in the  $p_T$ -evolution.

### 3.2.2 Multiple Parton Interactions in Pythia

The MPI, as said before, are also generated in a decreasing  $p_T$  sequence. So the hardest MPI is generated first. Than we can write the probability for an interaction,  $i$ , to occur at a scale  $p_T$  using a Sudakov-type expression:

$$\frac{d\mathcal{P}_{MPI}}{dp_T} = \frac{1}{\sigma_{ND}} \frac{d\sigma_{2 \rightarrow 2}}{dp_T} \exp \left( - \int_{p_T}^{p_T^{i-1}} \frac{1}{\sigma_{ND}} \frac{d\sigma_{2 \rightarrow 2}}{dp'_T} dp'_T \right) \quad (3.14)$$

### 3.2.3 Momentum and flavour conservation

One problem is to achieve momentum conservation, so we need to take into account the modification in the PDF by the  $i - 1$  interaction. To do that in `Pythia` the PDF are rescaled to the remaining available  $x$  range, adjusting their normalization.

We need to take into account the momentum fraction  $x_i$  removed from the hadron remnant by the  $i - th$  interaction. This is done evaluating PDF not at  $x_i$  but at a rescaled value

$$x'_i = \frac{x_i}{X} \quad \text{with:} \quad X = 1 - \sum_{j=1}^{i-1} x_j \quad (3.15)$$

So using these quantities, we can rewrite our PDFs as:

$$f_i(x, Q^2) \longrightarrow \frac{1}{X} f_0\left(\frac{x}{X}\right) \quad (3.16)$$

where  $f_0$  the original one-parton inclusive PDFs.

Now, requiring also the flavour conservation and taking into account the number of valence and/or sea quarks involved in the preceding MPI. We have now the full forms of the PDFs used for the  $i - th$  MPI:

$$f_i(x, Q^2) = \frac{N_{fv}}{N_{fv0}} \frac{1}{X} f_{v0}\left(\frac{x}{X}, Q^2\right) + \frac{a}{X} f_{s0}\left(\frac{x}{X}, Q^2\right) + \sum_j \frac{1}{X} f_{cj0}\left(\frac{x}{X}, Q^2\right) \quad (3.17)$$

$$g_i(x, Q^2) = \frac{a}{X} g_0\left(\frac{x}{X}, Q^2\right) \quad (3.18)$$

where:

- $f_i(x, Q^2)$  ( $g_i(x, Q^2)$ ) are the squeezed PDFs for quarks (gluons)
- $N_{fv}$  the number of remaining valence quarks of the given flavour
- $N_{fv0}$  the number of original valence quarks of the given flavour (for the proton we have  $N_u = 2$ ,  $N_d = 1$ )
- $f_{s0}$  the sea-quark PDF
- $f_{cj}$  the companion PDF, this arise from the splitting  $g \rightarrow q\bar{q}$  and a quark  $j$  is kicked out.

The factor  $a$  is defined to satisfy the total momentum sum rule.

### 3.2.4 Impact Parameter Dependence

The simplest hypothesis for the multiple interaction simulation, it is to assume the same initial state for all hadron collisions without dependencies on the impact parameter.

The more realistic scenario it is to include the possibility that each collision could be characterized by a different impact parameter  $b$ , where a small  $b$  value correspond to a large overlap between the two hadrons this is related to the probability of parton-parton interaction to take place.

It is necessary to make some assumption on the matter distribution inside the proton. A possibility is to assume a spherically symmetric distribution inside a hadron at rest  $\rho(\mathbf{x}) d^3x = \rho(r) d^3x$ . A Gaussian ansatz is the most simple choice but it appears to lead to a narrow multiplicity distribution and too little of a pedestal effect. So the choice is a double Gaussian:

$$\rho(r) \propto \frac{1 - \beta}{a_1^3} \exp\left\{-\frac{r^2}{a_1^2}\right\} + \frac{\beta}{a_2^3} \exp\left\{-\frac{r^2}{a_2^2}\right\} \quad (3.19)$$

where a fraction  $\beta$  of matter is contained in a radius  $a_2$ , and the rest in a larger radius  $a_1$ .

Now for a given matter distribution  $\rho(r)$ , the time-integrated overlap function of the incoming hadrons during the collision is given by:

$$\mathcal{O}(b) = \int dt \int d^3x \rho(x, y, z) \rho(x + b, y, z + t) \quad (3.20)$$

Assuming the matter distribution function in Eq. 3.19 and inserting it into Eq. 3.20 we get the following expression

$$\mathcal{O}(b) \propto \frac{(1 - \beta)^2}{2a_1^2} \exp \left\{ -\frac{b^2}{2a_1^2} \right\} + \frac{2\beta(1 - \beta)}{a_1^2 + a_2^2} \exp \left\{ -\frac{b^2}{a_1^2 + a_2^2} \right\} + \frac{\beta^2}{2a_2^2} \exp \left\{ -\frac{b^2}{2a_2^2} \right\} \quad (3.21)$$

this is useful to quantify the effect of overlapping protons. The larger is  $\mathcal{O}(b)$  the more probable are parton-parton scatters between the incoming protons  $\langle \tilde{n} \rangle \propto \mathcal{O}(b)$ .

So, these assumption change the so-far Poissonian nature of the framework<sup>1</sup>

$$\mathcal{P}(\tilde{n}) = \langle \tilde{n} \rangle^{\tilde{n}} \frac{e^{-\langle \tilde{n} \rangle}}{\tilde{n}!} \quad (3.22)$$

Now the request that at least one parton interactions in the hadron-hadron collision, ensures that we get a finite total cross section. So the probability that an event is produced by two hadrons scattering with impact parameter  $b$ :

$$\mathcal{P}_{\text{int}} = \sum_{\tilde{n}} = 1 - \mathcal{P}_{\tilde{n}(b)} = 1 - \mathcal{P}_0 = 1 - e^{-\langle \tilde{n}(b) \rangle} = 1 - e^{-k\mathcal{O}(b)} \quad (3.23)$$

So we have now that the number of interaction per event is give by (for impact parameter  $b$ ):

$$\langle n(b) \rangle = \frac{\langle \tilde{n}(b) \rangle}{\mathcal{P}_{\text{int}}} \quad (3.24)$$

so, this have modified the probability distribution of interactions number from a Poissonian to a narrower one at each  $b$  fixed.

### 3.2.5 Parton rescattering

It is not necessary that the partons undergoing to a MPI are a different partons couple from the one scattered before. As shown in Fig. 3.2 MPI can also arise when a parton scatters more than once against partons from the other beam, this is call *parton rescattering*.

We can have 3 type of MPI:

1. No ones of the partons that enter in the second scattering undergoes to another scatter before
2. Only one of the two partons have already been scattered
3. Both the partons have already been scattered before.

---

<sup>1</sup>the Poissonian distribution, in Eq. 3.22 describe the probability of having  $n$  interactions at each impact parameter. If the matter distribution have an infinite tail (like our in Eq. 3.19) events may be obtained with arbitrarily large  $b$  values. This can be a problem for the definition of the total hadron-hadron cross section

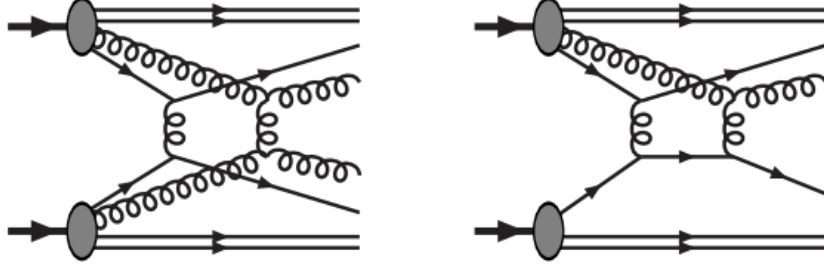


Figure 3.2: Add caption

The second and the third are the rescatters the overall influence of rescatters in proton-proton interactions was estimated to be small respect to the first case with distinct  $2 \rightarrow 2$  scatters.

The simulation of parton rescatters start from the evaluation of the parton density as:

$$f(x, Q^2) \longrightarrow \underbrace{f_{\text{rescaled}}(x, Q^2)}_{\text{hadron remnant}} + \underbrace{\sum_n \delta(x - x_n)}_{\text{scattered parton(s)}} \quad (3.25)$$

The  $\delta$  take into account the scattered partons that have a fixed momentum fraction  $x_n$ . While the hadron remnant is still described by a continuous momentum density, rescaled to achieve momentum conservation:

$$\int_0^1 \left( f_{\text{rescaled}}(x, Q^2) + \sum_n \delta(x - x_n) \right) dx = 1 \quad (3.26)$$

### 3.2.6 Interleaving of Multiple Interaction and Parton Shower

As discussed above the MPI are simulated in Pythia following a decreasing  $p_T$  evolution. So we have that the total probability is given from the composition of the various contributions:

$$\begin{aligned} \frac{d\mathcal{P}}{dp_T} = & \left( \frac{d\mathcal{P}_{MPI}}{dp_T} + \sum \frac{d\mathcal{P}_{ISR}}{dp_T} + \sum \frac{d\mathcal{P}_{FSR}}{dp_T} \right) \times \\ & \times \exp \left\{ - \int_{p_T}^{p_T^{i-1}} \left( \frac{d\mathcal{P}_{MPI}}{dp'_T} + \sum \frac{d\mathcal{P}_{ISR}}{dp'_T} + \sum \frac{d\mathcal{P}_{FSR}}{dp'_T} \right) dp'_T \right\} \end{aligned} \quad (3.27)$$

In Fig. 3.3 are shown 4 parton-parton interactions with their associated showers (ISR and FSR). The downward evolution correspond to read the graph from top to bottom. The first hard interaction occur at a scale  $p_T = p_{T1}$  while the following ones at lower scales  $p_{T2}$ ,  $p_{T3}$ ,  $p_{T4}$ . Each interaction is associated with its radiation the first one occurs at  $p_T = p'_{T1}$ . The scatterings that occur at  $p_{T2}$  and  $p_{T3}$  are originating from the same mother parton.

This diagram is related to one of the two hadrons. The full event can be illustrated if a similar diagram is drawn for the other hadron and connected to the full black circles.

### 3.2.7 Beam Beam Remnants and primordial $k_T$

What is left after that the evolution is end? the evolution in  $p_T$  can create an arbitrary complicate final state.

This final state contains contributions from the scattered and unscattered partons that don't enter the  $p_T$  evolution. The last are the so called Beam Beam Remnants (BBR). BBR take



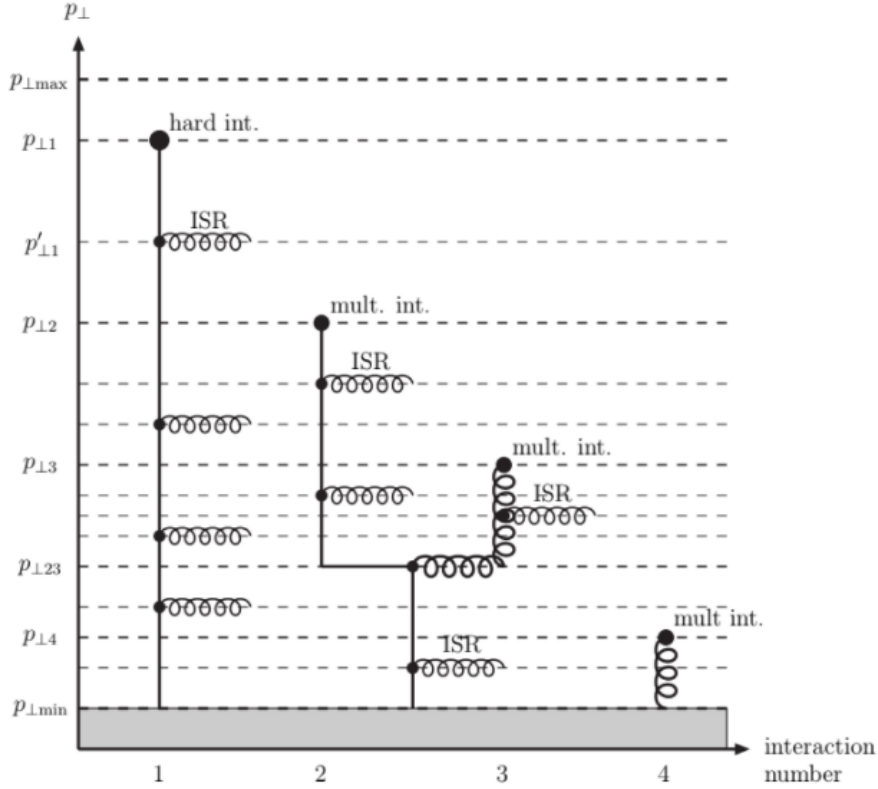


Figure 3.3: add caption

into account the number of valence quark remaining and the number of sea quark required for the overall flavour conservation.

To ensure momentum conservation the BBR have to take all the remaining longitudinal momentum that is not extracted by the MPI initiators.

### Primordial $k_T$

We have considered only the longitudinal momentum. in a real scenario partons inside the hadrons are fermions so are expected to have a non zero initial transverse momentum arising from the Fermi motion inside the incoming hadrons. This is denoted as "Primordial  $k_T$ " this is different from the transverse momentum derived from DGLAP shower evolution or from the hard interaction.

Based on Fermi motion alone, one would expect a value of the primordial  $k_T$  is estimate as:

$$k_T \simeq \frac{\hbar}{r_p} \approx \frac{0.2 \text{ GeV} \cdot \text{fm}}{0.7 \text{ fm}} \approx 0.3 \text{ GeV} \quad (3.28)$$

but for example to reproduce the data for the the  $p_T$  distributions of  $Z$  bosons produced in hadron-hadron collisions, one need a larger contribution. This phenomenon has not a satisfactory explanation. Until an explanation is found the idea is to consider a effective primordial  $k_T$  for the initiators larger than the one in Eq. 3.28.

In **Pythia** the primordial  $k_T$  is assigned to initiators sampling a Gaussian distributions for  $p_x$  and  $p_y$  independently with variable width  $\sigma$

$$\sigma(Q, \hat{m}) = \frac{Q_{1/2} \sigma_{\text{soft}} + Q \sigma_{\text{soft}}}{Q_{1/2} + Q} \frac{\hat{m}}{\hat{m}_{1/2} + \hat{m}} \quad (3.29)$$

Where  $Q$  is the hard-process renormalization scale for the main hard process and the  $p_T$  scale for subsequent MPI.  $\sigma_{\text{soft}}$ ,  $\sigma_{\text{hard}}$  are the minimum and maximum width,  $\hat{m}$  is the invariant mass, while  $Q_{1/2}$  and  $\hat{m}_{1/2}$  the halfway values between the two extremes.

### 3.2.8 Color Reconnection and Hadronization

The last important step at parton level is the color reconnection. Color reconnection is motivated by that MPI leads to different color strings. In the previous step the planar limit of the QCD was assumed where  $N_c \rightarrow \infty$ . Now, moving to real case where  $N_c = 3$  all this strings that can be overlapped in physical space can be reconnected. The basic idea is to reconnect strings in order to reduce the total string length; and thus the potential energy.

In PYTHIA8 the reconnection is performed giving to each system a probability to reconnect given by:

$$\mathcal{P}_{\text{rec}} = \frac{p_{T\text{rec}}^2}{(p_{T\text{rec}}^2 + p_T^2)} \quad p_{T\text{rec}} = R \times p_{T0} \quad (3.30)$$

the `ColorReconnection:range`  $R$  is a user-tunable parameter while  $p_{T0}$  is the same parameter defined in MPI simulation.

Whit this definition for the probability it is clear that system with low  $p_T$  are more likely that can reconnect to other, this idea find is justification in the fact that at low  $p_T$  value correspond a larger spatial extension and so these strings have more chance to overlap with other and so to reconnect.

The hadronization takes all these partons (color strings) and transform it in a set of color-neutral hadrons. Hadronization is based solely on the *Lund string fragmentation model* [23, 24].

Lund model basic idea is to break the color line and to reduce the total string length the string is representative of the potential

$$V(r) = -\frac{a}{r} + \kappa r \quad \text{with } \kappa \approx 1 \text{ GeV/fm} \quad (3.31)$$

where  $\kappa$  is the string tension. This potential is a combination of an attractive (Coulomb) potential ad a linear potential that phenomenologically include quarks confinement. The linear potential is the dominating part at increasing distance values, so the energy increase with distance.

The simplest case is the one in Fig. 3.4: The  $q\bar{q}$  system evolves in space increasing the string length at some point the distance is to large and is convenient to break the string into two strings.

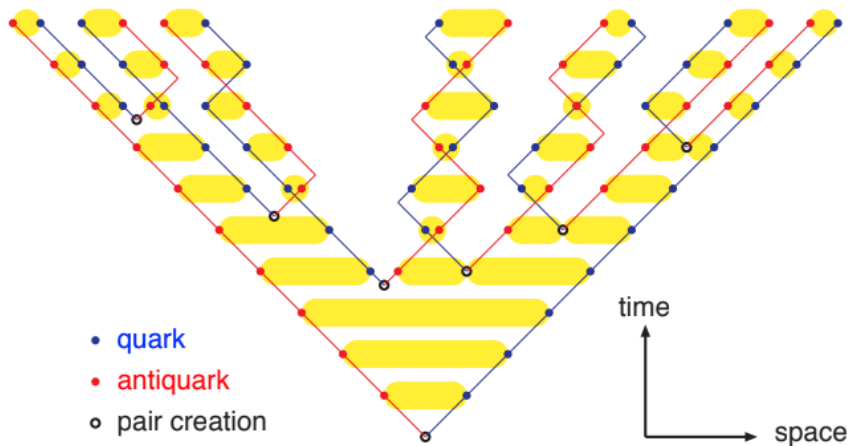


Figure 3.4: ADD

The hadronization step confines the quark into hadrons, than hadrons can undergoes to hadron rescattering and decay. These are the hadrons that are revealed by the detector.

### 3.3 PYTHIA summary

To summarize PYTHIA8 is able to simulate an high energy hadron-hadron collision. The evolution of the system is simulated in a common decreasing- $p_T$  sequence the master formula for the evolution is written in Eq. 3.27.

This evolution start from an hard scale that is the scale of the main parton hard scattering that is described by a fixed ME calculation, PYTHIA can be interfaced with external frameworks for the ME step as POWHEG and MAD-GRAPE5 AMC@NLO. Then the parton shower is started with the simulation of ISR, FSR and MPI also the parton rescattering is allowed. Once the evolution is ended the hadronization transform the partons in a set of final hadrons these hadron than decay and rescatters against each others before the detection.

All these processes are describe mainly by phenomenological model, due to the not-known-by-first-principles softQCD description. The use of these phenomenological introduces lot of free parameters (some are been pointed out in the previously sections) that have to be tuned with data to give to PYTHIA the ability of simulate real events.

In the next chapter we are going to focus on the study of the underlying event and so on the soft events related to an hard scattering.

# Chapter 4

## Observable to Study the Underlying Event

The UE are all the processes not associated with the primary hard scatter in an hadron-hadron collision.

All the process described in the previous section: ISR and FSR, MPI, and BBR and their interactions with color exchanges among them, contribute to the Underlying Event (UE) in the proton-proton collision.

The most of the observable to study the UE are sensible only to the sum of these contributions and not to the single ones so a good description of all this process and their interplay is really important.

### 4.1 Minimum Bias Measurements and Underlying Event topology

A Minimum Bias (MB) measurement is a collection of inelastic events with a loose event selection. The events are collected requiring the minimal interaction with the detector (the smallest possible bias). The most of the interaction in MB observation are soft,  $p_T \lesssim 2$  GeV. The study of the UE require at least one hard scattering ( $p_T \gtrsim 2$  GeV) presence.

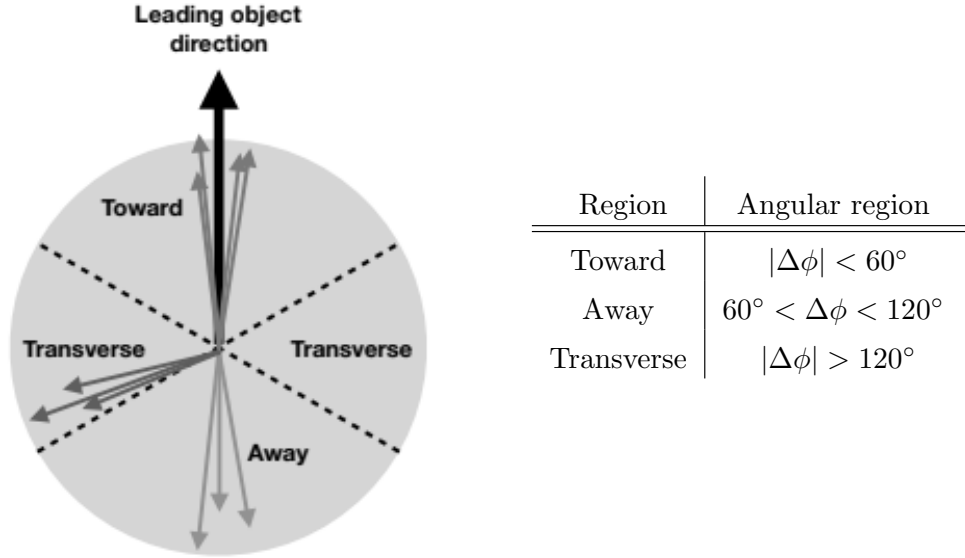
To study the UE the topological structure of an hadron hadron collision is used. On an event-by-event analysis the direction of the leading object is used to define regions in the  $\eta - \phi$  space. Where  $\eta$  is the pseudorapidity defined as  $\eta = -\log \tan\left(\frac{\theta}{2}\right)$ , while  $\phi$  is the azimuthal angle in the  $x - y$  plane. The last one is defined from the direction of the leading object as  $\Delta\phi = \phi - \phi_{\text{leading}}$ .

The regions classification is shown in Fig. 4.1, we have:

- **Toward region:** the region that contains the leading object, this region contains the most of the particle produced by the hard interaction.
- **Away region:** this region contains the objects that recoil against the leading object, also this region contains mostly the particles produced by the hard interaction.
- **Transverse regions:** the two transverse regions are the most sensitive to UE.

The transverse regions are also separated in:

- **TransMAX:** This is the transverse region that contains the *maximum* number of charged particles, or scalar  $p_T$  sum of charged particles. This regions includes both MPI and hard-process contamination.



**Figure 4.1:** This figure shows the four regions for the description of the UE on a event-by-event analysis. the angular values are shown in the table. The regions are defined starting from the leading object direction. The toward and away regions contain most of the contribution from the hard scattering (e.g. in a  $t\bar{t}$  production event the two quark  $t$  are located in these regions); while, the transverse region are the ones in between the twp other regions, these are the most important for the study of the UE.

- **TransMIN:** is the transverse region that contains the *minimum* number of charged particles, or scalar  $p_T$  sum of charged particles. This region is the most sensitive to MPI effects.

The leading object definition depend on the type of event under observation.

The charged-particle with largest  $p_T$  [25], the dilepton system in Drell-Yan observation [26, 27] or  $t\bar{t}$  events [28] can all be used as leading object in the analyses for the UE event.

So we are interest in studying the transMAX and transMIN region in particular the observable sensitive to UE in these regions are: the charged-particle density and the charged-particle scalar- $p_T$  sum density in the  $\eta - \phi$  space

In an hadron-hadron collision with two jets productions it is observed that also the transverse regions activity increase with energy, as shown in Fig. 4.2 This increment cannot be explained only by the increase of the FSR (away regions become broader increasing scale). To explain this rise one need to attribute activity in the transverse regions to MPI.

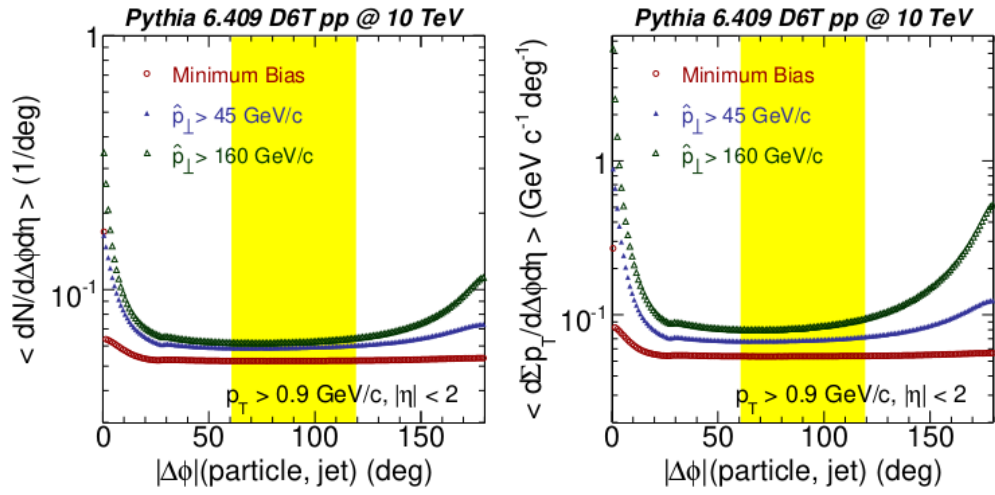


Figure 4.2: ADD



# Chapter 5

## Tune procedure, CP5 Tune and MCNNTUNES

The study of the underlying event, or more in general of softQCD processes require the use of Monte Carlo generator based on phenomenological model all these models introduce lot of free parameters that must be tuned on experimental data in order to obtain meaningful results. The procedure of estimate the best parameters values is called *tune*.

The tune procedure can be really computational expensive, can require to run the generator a very large amount of times. And usually these MC generator are really expansive in computation time required for a single job.

To tune some parameters the number of jobs you have to run increase with the number of parameter you want to tune. different approaches have been used in the past to tune these MC generator:

- 1) **By eye approach:**
- 2) **Brute force approach:**
- 3) **Parametrization-based approach:**





# Bibliography

- [1] Steven Weinberg. A model of leptons, Nov 1967. URL <https://link.aps.org/doi/10.1103/PhysRevLett.19.1264>.
- [2] Sidney D Drell and Tung-Mow Yan. Partons and their applications at high energies, 1971. ISSN 0003-4916. URL <https://www.sciencedirect.com/science/article/pii/0003491671900716>.
- [3] J M Campbell, J W Huston, and W J Stirling. Hard interactions of quarks and gluons: a primer for lhc physics, Dec 2006. ISSN 1361-6633. URL <http://dx.doi.org/10.1088/0034-4885/70/1/R02>.
- [4] L N Lipatov. The parton model and perturbation theory, 1975. URL <http://cds.cern.ch/record/400357>.
- [5] Vladimir Naumovich Gribov and L N Lipatov. Deep inelastic ep scattering in perturbation theory, 1972. URL <https://cds.cern.ch/record/427157>.
- [6] G. Altarelli and G. Parisi. Asymptotic freedom in parton language, 1977. ISSN 0550-3213. URL <https://www.sciencedirect.com/science/article/pii/0550321377903844>.
- [7] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and  $e^+ e^-$  Annihilation by Perturbation Theory in Quantum Chromodynamics., 1977.
- [8] W.J. Stirling. private communication. URL <http://www.hep.ph.ic.ac.uk/~wstirlin/plots/plots.html>.
- [9] F. Bloch and A. Nordsieck. Note on the radiation field of the electron, Jul 1937. URL <https://link.aps.org/doi/10.1103/PhysRev.52.54>.
- [10] Toichiro Kinoshita. Mass singularities of feynman amplitudes, 1962. URL <https://doi.org/10.1063/1.1724268>.
- [11] T. D. Lee and M. Nauenberg. Degenerate systems and mass singularities, Mar 1964. URL <https://link.aps.org/doi/10.1103/PhysRev.133.B1549>.
- [12] A. Kulesza, G. Stermann, and W. Vogelsang. Electroweak vector boson production in joint resummation, 2002. URL <https://arxiv.org/abs/hep-ph/0207148>.
- [13] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to pythia 8.2, Jun 2015. ISSN 0010-4655. URL <http://dx.doi.org/10.1016/j.cpc.2015.01.024>.

- [14] Manuel Bähr, Stefan Gieseke, Martyn A. Gigg, David Grellscheid, Keith Hamilton, Oluseyi Latunde-Dada, Simon Plätzer, Peter Richardson, Michael H. Seymour, Alexander Sherstnev, and Bryan R. Webber. Herwig++ physics and manual, Nov 2008. ISSN 1434-6052. URL <http://dx.doi.org/10.1140/epjc/s10052-008-0798-9>.
- [15] T Gleisberg, S Hoeche, F Krauss, A Schaelicke, S Schumann, and J Winter. Sherpa 1. , a proof-of-concept version, Feb 2004. ISSN 1029-8479. URL <http://dx.doi.org/10.1088/1126-6708/2004/02/056>.
- [16] E. Boos, M. Dobbs, W. Giele, I. Hinchliffe, J. Huston, V. Ilyin, J. Kanzaki, K. Kato, Y. Kurihara, L. Lonnblad, M. Mangano, S. Mrenna, F. Paige, E. Richter-Was, M. Seymour, T. Sjostrand, B. Webber, and D. Zeppenfeld. Generic user process interface for event generators, 2001.
- [17] Stefano Catani, Frank Krauss, Bryan R Webber, and Ralf Kuhn. Qcd matrix elements + parton showers. *Journal of High Energy Physics*, 2001(11):063–063, Nov 2001. ISSN 1029-8479. doi: 10.1088/1126-6708/2001/11/063. URL <http://dx.doi.org/10.1088/1126-6708/2001/11/063>.
- [18] Stefano Frixione and Bryan R Webber. Matching nlo qcd computations and parton shower simulations. *Journal of High Energy Physics*, 2002(06):029–029, Jun 2002. ISSN 1029-8479. doi: 10.1088/1126-6708/2002/06/029. URL <http://dx.doi.org/10.1088/1126-6708/2002/06/029>.
- [19] Stefano Frixione, Paolo Nason, and Bryan R Webber. Matching nlo qcd and parton showers in heavy flavour production. *Journal of High Energy Physics*, 2003(08):007–007, Aug 2003. ISSN 1029-8479. doi: 10.1088/1126-6708/2003/08/007. URL <http://dx.doi.org/10.1088/1126-6708/2003/08/007>.
- [20] Stefano Frixione and Bryan R. Webber. The mc@nlo 3.1 event generator, 2005.
- [21] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Luigi Del Debbio, Stefano Forte, Patrick Groth-Merrild, Alberto Guffanti, Nathan P. Hartland, Zahari Kassabov, José I. Latorre, Emanuele R. Nocera, Juan Rojo, Luca Rottoli, Emma Slade, and Maria Ubiali. Parton distributions from high-precision collider data. *The European Physical Journal C*, 77(10), Oct 2017. ISSN 1434-6052. doi: 10.1140/epjc/s10052-017-5199-5. URL <http://dx.doi.org/10.1140/epjc/s10052-017-5199-5>.
- [22] FRANK SIEGERT. Monte-carlo event generation for the lh, 2010. URL <http://etheses.dur.ac.uk/484/>.
- [23] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand. Parton fragmentation and string dynamics. *Physics Reports*, 97(2):31–145, 1983. ISSN 0370-1573. doi: [https://doi.org/10.1016/0370-1573\(83\)90080-7](https://doi.org/10.1016/0370-1573(83)90080-7). URL <https://www.sciencedirect.com/science/article/pii/0370157383900807>.
- [24] Torbjorn Sjostrand. Jet Fragmentation of Nearby Partons. *Nucl. Phys. B*, 248:469–502, 1984. doi: 10.1016/0550-3213(84)90607-2.
- [25] Underlying Event Measurements with Leading Particles and Jets in pp collisions at  $\sqrt{s} = 13$  TeV. Technical report, CERN, Geneva, 2015. URL <https://cds.cern.ch/record/2104473>.

- [26] Serguei Chatrchyan et al. Measurement of the underlying event in the Drell-Yan process in proton-proton collisions at  $\sqrt{s} = 7$  TeV. *Eur. Phys. J. C*, 72:2080, 2012. doi: 10.1140/epjc/s10052-012-2080-4.
- [27] A. M. Sirunyan et al. Measurement of the underlying event activity in inclusive Z boson production in proton-proton collisions at  $\sqrt{s} = 13$  TeV. *JHEP*, 07:032, 2018. doi: 10.1007/JHEP07(2018)032.
- [28] Albert M. Sirunyan et al. Study of the underlying event in top quark pair production in pp collisions at 13 TeV. *Eur. Phys. J. C*, 79(2):123, 2019. doi: 10.1140/epjc/s10052-019-6620-z.