# Multiple Parton Interactions in Photoproduction at HERA/H1

## Dissertation

zur Erlangung des Doktorgrades
des Department Physik
der Universität Hamburg

vorgelegt von

Lluís Martí Magro

aus Barcelona

Hamburg
2008

| | |
|---|---|
| Gutachter/Gutachterin der Dissertation: | Prof. Dr. E. Elsen<br>Dr. H. Jung |
| Gutachter/Gutachterin der Disputation: | Prof. Dr. J. Haller<br>Dr. H. Jung |
| Datum der Disputation: | 22.01.2009 |
| Vorsitzende/Vorsitzender des Prüfungsausschusses: | Prof. Dr. C. Hagner |
| Vorsitzender des Promotionsausschusses: | Prof. Dr. R. Klanner |
| MIN-Dekan des Departaments Physik der Fakultät: | Prof. Dr. Arno Frühwald |

# Abstract

Photoproduction data of HERA-I are analysed by requiring dijets with transverse momenta of at least 5 GeV. The two jets define in azimuth a towards region (leading jet), an away region (usually the 2nd jet) and transverse regions between them. The charged particle and jet with low transverse momentum multiplicity, so called minijets, are measured in these regions as a function of the variables $x_\gamma^{obs}$ and $P_T^{Jet_1}$ (leading jet). The measurement is compared to predictions including parton showers and matrix elements at leading order in $\alpha_s$. Some predictions include contributions from multiple parton interactions and use different parton evolution equations. It was found that existing MC programs do not fully describe the measurements but the description can be improved by including multiple parton interactions.



# Kurzfassung

HERA-I Photoproduktions-Daten werden analysiert, wobei mindestens zwei Jets mit Transversalimpuls größer als 5 GeV verlangt werden. Die beiden Jets definieren die folgenden azimutal Regionen: die Towards- (Richtung des führenden Jets), Away- (üblicherweise in Richtung des zweiten Jets) und transversalen Regionen zwischen ihnen. Die Multiplizität von geladenen Teilchen und Jets mit kleinem Transversalimpuls, so gennante Minijets, werden in diesen Regionen als Funktion der Variablen $x_\gamma^{obs}$ und $P_T^{Jet_1}$ (führender Jet) gemessen. Die Messungen werden mit Vorhersagen verglichen, die Parton Schauer und Matrix Elemente in führender Ordung von $\alpha_s$ enthalten. Manche Vorhersagen beinhalten Beiträge von vielfachen Parton-Parton-Wechselwirkungen und benutzen andere Parton-Entwicklungsgleichungen. Es wurde herausgefunden, dass die existierenden MC Programme die Messungen nicht vollständig beschreiben können. Die Beschreibung der Daten ist besser, wenn mehrfache Parton-Wechselwirkungen berücksichtigt werden.

# Contents

# Introduction

The kind of questions elementary particle physics asks are: What is matter made of? Which are the forces acting in nature? What are the properties of matter and the forces?

At the subatomic level and at even lower scales only three forces are considered: the strong, the weak and the electromagnetic forces. The particles currently considered as elementary particles are the leptons (electrons, muons, tauons and their respective neutrinos and their antiparticles), quarks (up, down, strange, charm, bottom and top and their respective antiparticles) and finally the gauge bosons which are the mediators of the forces between the particles (photons, $W^{\pm}$, $Z^{\circ}$ and the gluons).

What has been considered to be the fundamental particles and forces in physics has been changing throughout history. The aim has always been to obtain fundamental knowledge about the elementary constituents through unifying phenomena which initially looked different. An outstanding example is the unification of electricity and magnetism in electromagnetism done by Maxwell. Protons were considered in the past as elementary particles but now we believe they consist on quarks and gluons.

Gluons are the bosons responsible of the strong force. Similarly to the electromagnetic force, gluons are massless as photons are. Colour was introduced to account for an internal degree of freedom of quarks. Colour is so to say the strong force 'charge'. The most important difference between the electromagnetic and the strong forces is that gluons carry colour. This means that gluons interact with themself. This is the beauty and the difficulty of Quantum Chromodynamics (QCD), the theory that describes the strong force.

Since protons are made of quarks and gluons, the strong force is prominent inside them. At HERA where electrons and protons were collided, one subject of investigation was the understanding of the proton content. The idea is to use an elementary particle, i.e. an electron, to study a composed particle like the proton. The electron and the proton interact via a photon or a $Z^{\circ}$ particle in neutral current events or via a $W^{\pm}$ particle in charged current events.

In this analysis, the virtuality of the exchanged boson is very low and it is enough to consider only photon exchange. Given the low virtuality, the exchanged photon lives long enough and is able to split into a quark-antiquark and form a hadronic structure. In this case we have two hadronic particles colliding and as in hadron colliders multiple particle

interactions are possible (MPI).

The aim of this analysis is the study of MPI. Two different observables are defined: charged particle and low transverse momentum jet average multiplicity. These observables are measured at HERA and compared to theoretical predictions.

The thesis starts with a theoretical overview. It introduces concepts needed to understand the physics of $ep$ collisions and is a preparation to the subject of muliple parton interactions in the next chapter. Chapter 3 describes the most relevant detector components at H1 for this analysis. The next two chapters describe the event selection and the procedure to correct the data for detector effects. Chapter 6 presents the measurement of the two observables and compares them with theoretical predictions.

# Chapter 1

# Theoretical Overview

In this chapter a theoretical overview of the physics at HERA is given. In this analysis QCD is of special interest and therefore this introduction will be focused on the QCD aspects of the HERA physics in the photoproduction regime where the exchanged boson is a photon.

## 1.1 ep Scattering and Kinematics

At HERA electron[1] and proton beams were accelerated at 27.6 GeV and 920 GeV energies, respectively. The mediator boson was a virtual photon, $\gamma^*$, or a $Z^\circ$ in neutral current scattering or a $W^\pm$ in charged current events. Electron-proton scattering is depicted in figure 1.1.

After the boson exchange the proton can remain intact or break up, elastic or inelastic scattering, respectively. The four momenta of the particles are:

- Incoming electron: $k = (E_e, \vec{k})$
- Outgoing electron: $k' = (E_{e'}, \vec{k'})$
- Incoming proton: $p = (E_p, \vec{p})$
- Boson mediator: $q = k - k'$

Of special importance in the description of the $ep$ scattering are the following variables:

- Centre-of-mass energy[2]: $s = (k + p)^2 \simeq 4E_eE_p$

[1] Also positrons were accelerated at the same energy. In general, we will refer to both just as electrons

[2] At HERA $\sqrt{s} \simeq 318$ GeV.

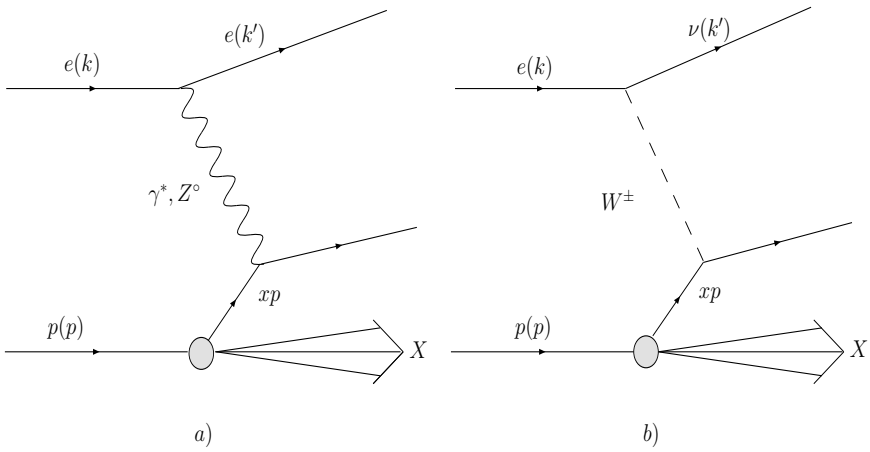


**Figure 1.1:** *ep* scattering neutral current a) with $\gamma^*$, or a $Z^\circ$ as the boson mediator and charged current b) with $W^\pm$ as the boson mediator.

- Bjorken variables:
  $x = \frac{Q^2}{2\ p\cdot q}$
  $y = \frac{p\cdot q}{p\cdot k}$
- Boson virtuality: $Q^2 = -q^2 = -(k-k')^2 = x\ y\ s$
- Boson-proton centre-of-mass energy: $W^2 = (q+p)^2 \simeq y\ s - Q^2$

The Bjorken variables take values between 0 and 1. In the quark parton model, $x$ is the fraction of the proton's momentum carried by the struck quark and $y$, the *inelasticity*, is the fraction of the incident electron's momentum carried by the exchanged boson in the proton rest frame. From the above definitions it is easy to see that for fixed beam energies only two variables are independent.

Roughly speaking the electron "illuminates" the proton content. In *ep* scattering two regimes are defined on the basis of the boson virtuality. For events with $Q^2 < 1$ GeV$^2$ the boson is almost on mass shell. In this regime only photon exchange will be considered since weak interactions are negligible due to the large masses of $Z^\circ$ and $W^\pm$. Since the photon is almost real, this regime is called *photoproduction*. At $Q^2 \gtrsim 1$ GeV$^2$, the virtuality provides a resolution scale and since the size of the proton is $\mathcal{O}(10^{-15}m)$, i.e. $Q^2 \approx 1$ GeV$^2$, in this regime it allows to prove the proton structure. This regime is called *deep inelastic scattering* (DIS)[3].

In this analysis, the photon virtuality is required to be in the range $Q^2 < 0.01$ GeV$^2$. At these virtuality values the proton may also break up. In this case we also have inelastic

[3]Here, 'deep' refers to the large virtuality values considered, $Q^2 > 1$ GeV$^2$, while 'inelastic' refers to the fact that the proton breaks up, $W^2 \gg m_p^2$, where $m_p$ is the mass of the proton.

scattering but the resolution scale is different to $Q^2$.

## 1.2 The Quark Parton Model

In the Quark Parton Model (QPM) the proton is seen as an incoherent sum of spin-$\frac{1}{2}$ point-like particles, called *partons* [1]. These partons are the valence quarks required by the SU(3) symmetry. Therefore, in QPM the fundamental process is the elastic electromagnetic interaction between two spin-$\frac{1}{2}$ point-like particles, $e + q \rightarrow e + q$. The cross section is proportional to the two leptonic tensors:

$$d\sigma \propto L_e^{\mu\nu} L_{\mu\nu}^q, \tag{1.1}$$

which leads to:

$$\frac{d\hat{\sigma}}{dy} = \frac{2\pi\alpha^2}{Q^4}\left[1 + (1-y)^2\right] \hat{s}\ e_{q_i}^2, \tag{1.2}$$

where $e_{q_i}$ is the charge of the struck quark and $\hat{s}$ is the quark-electron centre-of-mass energy ($\hat{s} = x\ s$). Equation 1.2 refers only to the quark-electron scattering. Now, we introduce $q_i(x)$ which is the probability that the struck quark $i$ carries a fraction $x$ of the proton's momentum.

The double differential cross-section for $e + p \rightarrow e + X$ is obtained when summing over all quarks in equation 1.2:

$$\begin{aligned}\frac{d^2\sigma}{dx\ dy} &= \sum_i \int_0^1 \frac{d\hat{\sigma}}{dy} x' q_i(x') \delta(x - x')\ dx' \\ &= \frac{2\pi\alpha^2}{Q^4}\left[1 + (1-y)^2\right]\ s\ \sum_i e_{q_i}^2\ x\ q_i(x)\end{aligned}$$

and similarly

$$\frac{d^2\sigma}{dx\ dQ^2} = \frac{2\pi\alpha^2}{xQ^4}\left[1 + (1-y)^2\right]\ \sum_i e_{q_i}^2\ x\ q_i(x). \tag{1.3}$$

We could do the same but now replacing the quark leptonic tensor $L_{\mu\nu}^q$ in equation 1.1 for a general hadronic tensor which parametrizes the hadronic current. The resulting differential cross section is:

$$\frac{d^2\sigma}{dx\ dy} = \frac{4\pi\alpha^2 s}{Q^4}\left[xy^2 F_1(x,y) + (1-y) F_2(x,y)\right] \tag{1.4}$$

$$\frac{d^2\sigma}{dx\ dQ^2} = \frac{2\pi\alpha^2}{xQ^4}\left[\left(1+(1-y)^2\right)F_2(x,Q^2) - y^2F_L(x,Q^2)\right], \tag{1.5}$$

where we introduce the *structure functions* $F_1$ and $F_2$ and $F_L = F_2 - 2xF_1$. Comparing equations 1.3 and 1.5 QPM predicts Bjorken scaling:

$$F_2(x,Q^2) = \sum_i e_{q_i}^2 x q_i(x), \tag{1.6}$$

i.e $F_2$ does depend on $x$ but not on $Q^2$, and:

$$F_L(x,Q^2) = 0 \Leftrightarrow 2xF_1 = F_2 \tag{1.7}$$

## 1.3 Quantum Chromodynamics

The QPM model considered only quarks. Since the quarks inside the proton by themself would not be stable inside the proton there should be some other particles keeping them together, or 'gluing' the quarks. The name gluon was given to the field quantum of strong interactions. Besides this problem, in the QPM partons were essentially free which contrasts with the fact that no free parton has been observed and only mesons and baryons are produced in all scattering experiments.

These two problems were solved by the theory of the strong interactions which in contrast to QED is a non-Abelian gauge theory, Quantum Chromodynamics (QCD)[4]. In this theory partons, quarks and gluons, carry a quantum number called *colour*. There are three colours and they are usually denoted by the colours red, green and blue. 't Hooft proved that the theory is renormalizable in 1971.

Two other major developments are *asymptotic freedom*[5] and *confinement*. Basically, when two colour connected partons are close together the strong force is weak (asymptotic freedom) while when moving apart the force becomes stronger (confinement). The asymptotic freedom tells us that at short distances, i.e. high energies, the partons in the proton can be seen as almost free particles as supposed for the quarks in the QPM while at long distances the interaction is very strong.

### 1.3.1 QCD Improved Parton Model

In the QPM only QED is involved. Therefore, the QPM can be seen as the zeroth order of QCD ($\mathcal{O}(\alpha_s^0)$). There are several processes to be added to QPM. Two of them will

[4] This was in the early 70s by Fritzsch, Gell-Mann and Leutwyler among others.
[5] Also in the early 70s by Politzer, Gross and Wilczek and were awarded the Nobel Prize in 2004.

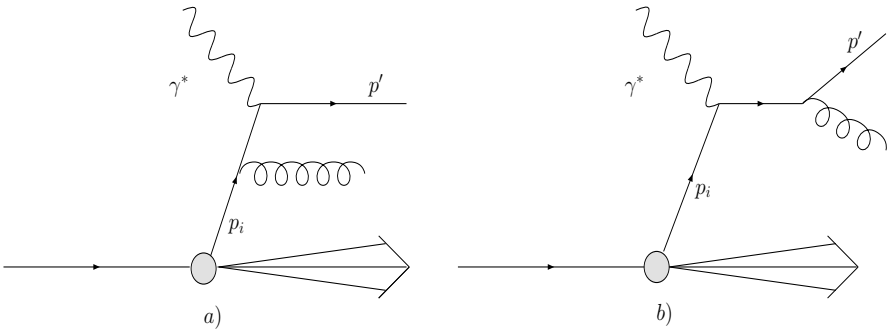


**Figure 1.2:** Gluon radiation from quarks via QCDC.

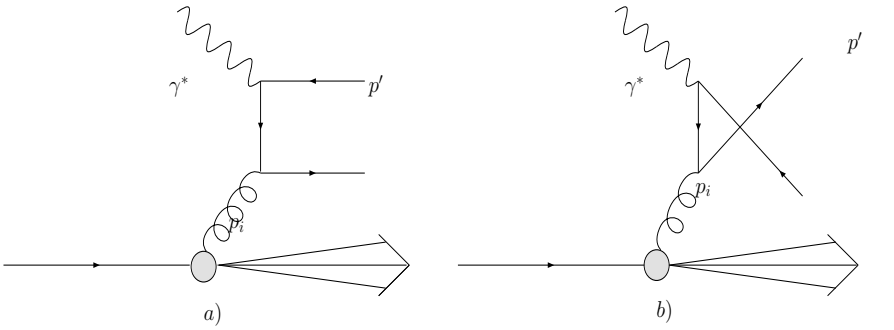


**Figure 1.3:** Boson gluon fusion (BGF) diagrams with $q\bar{q}$ production.

be considered first: gluon radiation from quark lines, or gluon Bremsstrahlung (QCD Compton, QCDC), and $g \to q\bar{q}$ (boson gluon fusion, BGF), shown in figures 1.2 and 1.3, respectively.

The QPM $\gamma^* p$, the cross section can be written as:

$$\begin{aligned}\sigma(x,Q^2) &= \sum_i \int_0^1 dz \int_0^1 d\xi q_i(\xi)\delta(x-z\xi)\underbrace{\hat{\sigma}(z,Q^2)}_{e_{q_i}^2\hat{\sigma}_\circ\delta(1-z)} \\ &= \hat{\sigma}_\circ \sum_i \int_x^1 \frac{d\xi}{\xi} q_i(\xi) e_{q_i}\delta(1-x/\xi) = \hat{\sigma}_\circ \sum_i e_{q_i}^2 q_i(x) = \hat{\sigma}_\circ \frac{F_2(x,Q^2)}{x}, \quad (1.8)\end{aligned}$$

where $\hat{\sigma}_\circ$ is the QPM parton cross section, $z = Q^2/(2p_i q) = x/\xi$ the analogue of the Bjorken $x$ for $\gamma^*$-parton scattering and $F_2(x,Q^2) = F_2(x)$, Bjorken scaling. It will seen that, by adding new contributions, BGF and QCDC for example, the Bjorken scaling is broken and $F_2(x) \to F_2(x,Q^2)$.

The QCDC partonic cross section can be written as a function of $\hat{\sigma}_\circ$ as:

$$\hat{\sigma}_{QCDC}(z) = \hat{\sigma}_\circ e_i^2 \frac{\alpha_s}{2\pi}\left[P_{qq}(z)\ ln\left(\frac{Q^2}{\kappa^2}\right) + C(z)\right], \text{where} \tag{1.9}$$

$$P_{qq}(z) = \frac{4}{3}\left[\frac{1+z^2}{1-z}\right] \tag{1.10}$$

is the probability for the $q \to q(z)g(1-z)$ splitting and is called *splitting function*. Here, $\kappa$ is introduced to regulate singularities in the integrals for low gluon (or quark) transverse momentum ($k \to 0$) in the parton CMS, *infrared divergency*.

Now, adding 1.9 to $\hat{\sigma}_\circ$ in 1.8

$$\frac{F_2(x,Q^2)}{x} = e_{q_i}^2 \int_x^1 \frac{d\xi}{\xi}\hat{\sigma}_\circ\ q_i(\xi,\mu^2)\left[\delta(1-\frac{x}{\xi}) + \frac{\alpha_s}{2\pi}P_{qq}\left(\frac{x}{\xi}\right)\ ln\left(\frac{Q^2}{\mu^2}\right) + \frac{\alpha_s}{2\pi}C(\frac{x}{\xi})\right] \tag{1.11}$$

where a new scale $\mu^2 \gg \kappa^2$ was introduced[6]. This means that the soft non-perturbative physics, i.e. long distances, are absorbed in the renormalized parton density $q_i(x,\mu^2)$. Now, the parton density $q_i(x) \to q_i(x,\mu^2)$:

$$q_i(x,\mu^2) = q_i^0(x) + \frac{\alpha_s}{2\pi}\int_x^1 \frac{d\xi}{\xi} q_i^0(\xi) P_{qq}\left(\frac{x}{\xi}\right)\ ln\left(\frac{\mu^2}{\kappa^2}\right). \tag{1.12}$$

Parton densities cannot be calculated perturbatively but only its variation with $ln\ \mu^2$:

$$\frac{\partial q_i(x,\mu^2)}{\partial ln\ \mu^2} = \frac{\alpha_s}{2\pi}\int_x^1 \frac{d\xi}{\xi} q_i(\xi,\mu^2) P_{qq}(x/\xi). \tag{1.13}$$

The BGF partonic cross section:

$$\hat{\sigma}_{BGF}(z) \simeq \hat{\sigma}_\circ e_i^2 \frac{\alpha_s}{2\pi}\left[P_{qg}(z)\ ln\left(\frac{Q^2}{\kappa^2}\right) + C(z)\right], \text{where} \tag{1.14}$$

$$P_{qg}(z) = \frac{1}{2}\left[z^2 + (1-z)^2\right] \tag{1.15}$$

is the $g \to q(z)\bar{q}(1-z)$ splitting function. If $\hat{\sigma}_{BGF}$ is convoluted with a gluon density $g$ and this is added to equation 1.11, equation 1.13 becomes:

[6] This is the collinear factorization scale, the scale at which the singularity is absorbed into the parton density. It was introduced by replacing $ln(Q^2/\kappa^2) \to ln(Q^2/\mu^2) + ln(\mu^2/\kappa^2)$.

$$\frac{\partial q_i(x,\mu^2)}{\partial ln\ \mu^2} = \frac{\alpha_s}{2\pi}\int_x^1 \frac{d\xi}{\xi}\left[q_i(\xi,\mu^2)P_{qq}(x/\xi) + g(\xi,\mu^2)P_{qg}(x/\xi)\right]. \tag{1.16}$$

Similar to equation 1.16 for the quark density, the gluon density evolves as:

$$\frac{\partial g_i(x,\mu^2)}{\partial ln\ \mu^2} = \frac{\alpha_s}{2\pi}\int_x^1 \frac{d\xi}{\xi}\left[\sum_i q_i(\xi,\mu^2)P_{gq}(x/\xi) + g(\xi,\mu^2)P_{gg}(x/\xi)\right], \text{ where} \tag{1.17}$$

$$P_{gq}(z) = \frac{4}{3}\left[\frac{1+(1-z)^2}{z}\right] \text{ and } P_{gg}(z) = 6\left[\frac{1-z}{z} + \frac{z}{1-z} + z(1-z)\right]. \tag{1.18}$$

Equations 1.16 and 1.18 are the DGLAP[7] *equations*. Although parton density functions cannot be derived from first principles they provide the description about how do parton density functions evolve from a scale to another one. This is achieved through the splitting functions. The structure function $F_2$, figure 1.4, was measured at HERA with a high precision. It represents an important measurement for other experiments since by the *factorization theorem* it is a universal function (section 1.3.3).

### 1.3.2 Infrared and Ultraviolet Divergences

In section 1.3.1 it was shown that infrared singularities appear in the cross section calculation. All these singularities were included in the parton distributions. By doing so the total cross section is factorized into short distance contributions on one side and all long distance contributions on the parton distributions. The scale which separates both is $\mu$, the *factorization scale*.

There are also *ultraviolet singularities*. These come from diagrams including loops where the integration over the momentum is unrestricted since they are virtual states in which the energy is not conserved. The technique to regularize them is called *renormalization* and introduces a new scale $\mu_R$, the *renormalization scale*. There are different schemes for doing it and the results depend on the choice. However, measured cross sections should be independent of the scale. Overviews on renormalization can be found in [2, 3].

### 1.3.3 Factorization Theorem in e$p$

In the QPM we made use of the assumption that the cross section can be factorized. The total cross section was expressed as the convolution of the probability of finding a parton

[7] These equations were develop by Dokshitzer, Gribov, Lipatov, Altarelli and Parisi in the late 70s and are known as DGLAP equations.

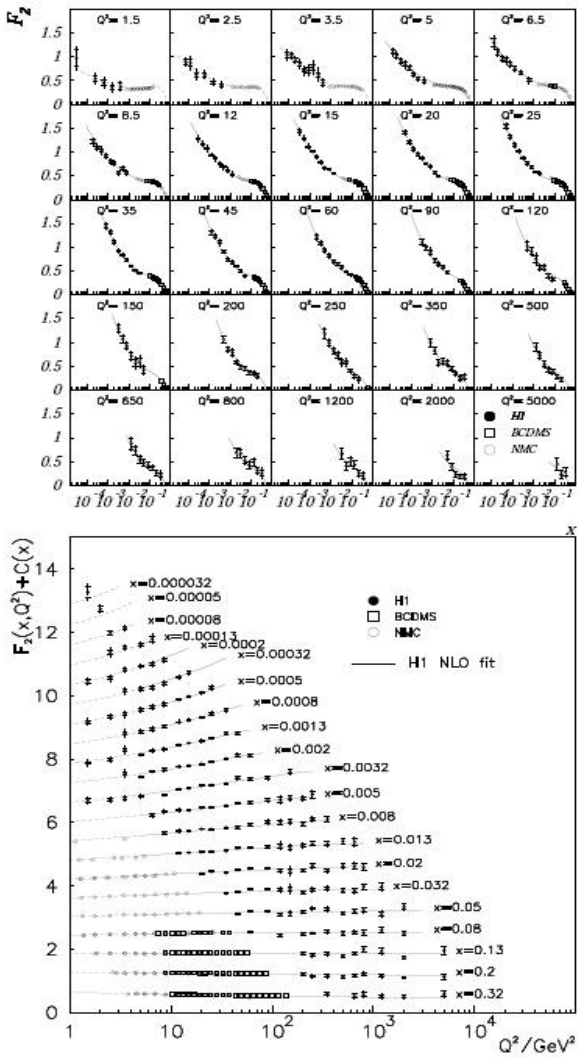


**Figure 1.4:** The structure function $F_2$ in $Q^2$ bins as a function of $x$ (top) and as a function of $Q^2$ for different $x$ values.

and the probability of a hard scattering. Therefore, the cross section in equation 1.8 is the convolution of $q_i$ with $\hat{\sigma}$.

In the improved QPM, infrared singularities appear and in order to avoid singularities the scale $\mu$ was introduced to keep long and short distance contributions factorized. Then, in order to avoid ultraviolet divergences another scale $\mu_R$ is introduced. The cross section can be written as:

$$\sigma = \sum_i \int_0^1 f_{i/p}(x, \mu^2) \hat{\sigma}_i \left(\hat{s}, \mu, \mu_R\right) \; dx, \tag{1.19}$$

where $f_{i/p}$ is the parton density function (quarks and gluons) of the parton $i$ in the proton which is universal and process independent. A factorization theorem for the DIS case can be found in [4]. For other processes like Drell Yan in hadron hadron collisions a factorization theorem has been proven but this is not the case for all known processes. An example of this is diffraction where a factorization theorem is available only at large $Q^2$. In the photoproduction regime, low $Q^2$, and in hadron hadron collisions, a factorization theorem for diffraction has not been proven.

### 1.3.4 Hadronization

Due to confinement, after the hard scattering which takes place at short distances, colourless particles are created, i.e hadrons. This process is called *hadronization*. Hadronization is dominated by long distance interactions which means that QCD becomes strong interacting and no perturbative calculations can be done. Hadronization is the combination of *fragmentation*, i.e colourless particle production from coloured partons, and the subsequent decay of the unstable hadrons into other hadrons.

Although fragmentation cannot yet be calculated from first principles, there are phenomenological models to describe this process. One of the most successful model is the *Lund string model* [5, 6] implemented in the JETSET Monte Carlo package [7]. Another model for fragmentation is the *cluster fragmentation model* [8] implemented in the Monte Carlo HERWIG [9].

#### 1.3.4.1 The Lund string model

The starting assumption in the string model is linear confinement between two colour connected partons moving apart. This means we can approximate the potential between

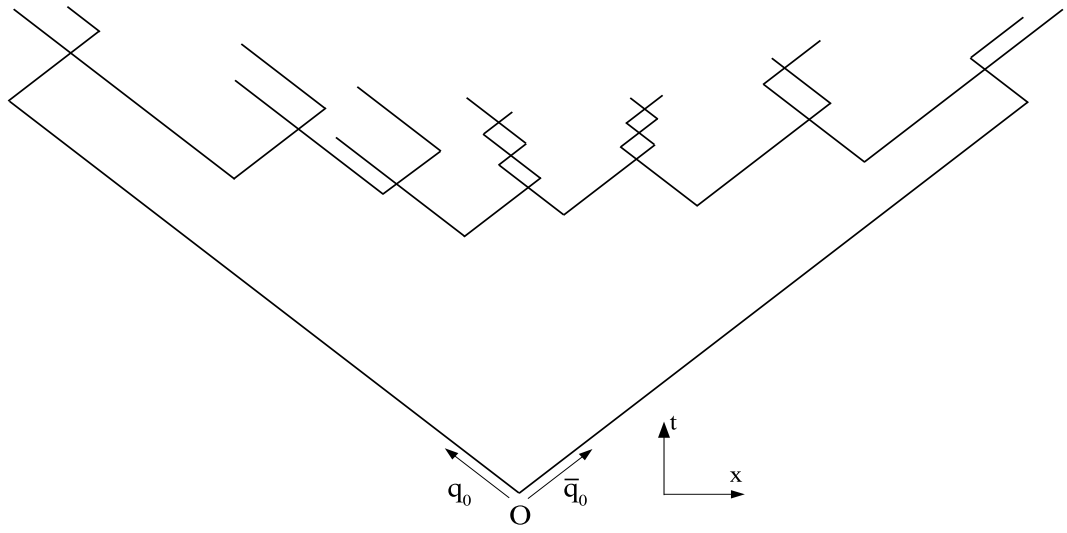


**Figure 1.5:** String fragmentation in t-x space.

the colour charges by the linear term[8]

$$V(r) = -\frac{4\alpha}{3r} + \kappa r \simeq \kappa r,$$

with $\kappa \approx 1$ GeV/fm and where $r$ is the distance between the charges. Thus, the energy stored in the string increases linearly. The picture of the process is that of a uniform colour flux tube between the colour charges. The model uses the dynamics of a massless relativistic string in order to get a Lorenz covariant and causal description of the energy flow.

The colour charges, while moving apart of each other from the origin $O$ (see figure 1.5), store energy in the string. This energy can be used to produce new pairs of charges and thereby breaking the string. The model predicts this to have a production rate of $exp\left(-\frac{\pi\mu_\perp^2}{\kappa}\right)$ calculated from quantum mechanical tunnelling, where $\mu$ is the mass of the charge and $\mu_\perp$ its transverse mass. This process stops when there is no more energy left in the string to produce new pairs.

The model predicts a suppression of heavy quark production, u : d : s : c $\approx 1 : 1 : 0.3 : 10^{-11}$. Therefore, charm quarks are expected to be produced only in the perturbative parton shower branchings $g \to q\bar{q}$ and not during the hadronization.

Since the different string breaks are causally disconnected, the order of the string breaks is arbitrary. Thus, the breaks can be chosen in any convenient order and starting from left to right or vice versa should be equivalent. This symmetry constrains the fragmentation

[8]This picture is supported [6] by Regge phenomenology, "onium" charmonium and bottonium spectroscopy [10], bag model calculations and lattice QCD results.

function $f(z)$, where $z$ is the fraction of the remaining momentum taken by the new particle resulting in the *Lund symmetric fragmentation function*:

$$f(z) \propto z^{-1}\ (1-z)^{a}\ exp\left(-bm_{\perp}^{2}/z\right),$$

where $a$ and $b$ are parameters of the model and $m_t$ is the hadrons transverse mass.

#### 1.3.4.2 The cluster fragmentation model

The cluster fragmentation model of hadronization divides fragmentation in two phases. The first one occurs at the end of the perturbative phase of parton showers after the hard scattering where clusters are formed. The second one is in the non-perturbative regime where hadrons arise from the decay of the clusters.

After the hard scattering, all partons evolve to lower scales. In this process, they can further branch into new partons until a given scale is reached, typically ∼1 GeV. At this scale non-perturbative effects dominate and *preconfinement* occurs. Preconfinement is the tendency of partons to be arranged in colour-singlet clusters. This is assumed to occur through $g \rightarrow q\bar{q}$ splittings (figure 1.6) where neighbour quarks from different gluons combine into clusters (represented by blobs) which can later further decay into new clusters of lower masses. Hadrons (black dots) are produced from these clusters.

The cluster mass distribution is universal and is assumed to represent the primordial resonances at initial stages of confinement. These clusters may decay isotropically in their rest frame into other clusters of lower masses via a two-body decay and these in turn into hadrons. The decay product into hadrons is chosen from a list of possible resonances with a probability which depends on the number of possible states (degeneracy of states for a given resonance) and the available phase space.

## 1.4 Photoproduction

At HERA the photoproduction regime, i.e. $Q^2 < 1$ GeV$^2$, dominates since, the cross section is proportional to $1/Q^4$ (equation 1.5). Experimentally, photoproduction is defined when the electron escapes detection or when the electron is tagged at very low scattering angles, as in this analysis where $Q^2 < 0.01$ GeV$^2$. In this case the electron can be seen as a source of quasi-real photons. The energy of the exchanged photon depends on the inelasticity, $y$, the energy fraction of the electron carried by the photon.

In the *equivalent photon approximation* [11,12] the photon flux of a fast moving electron is considered to be similar with the electromagnetic field with an energy distribution

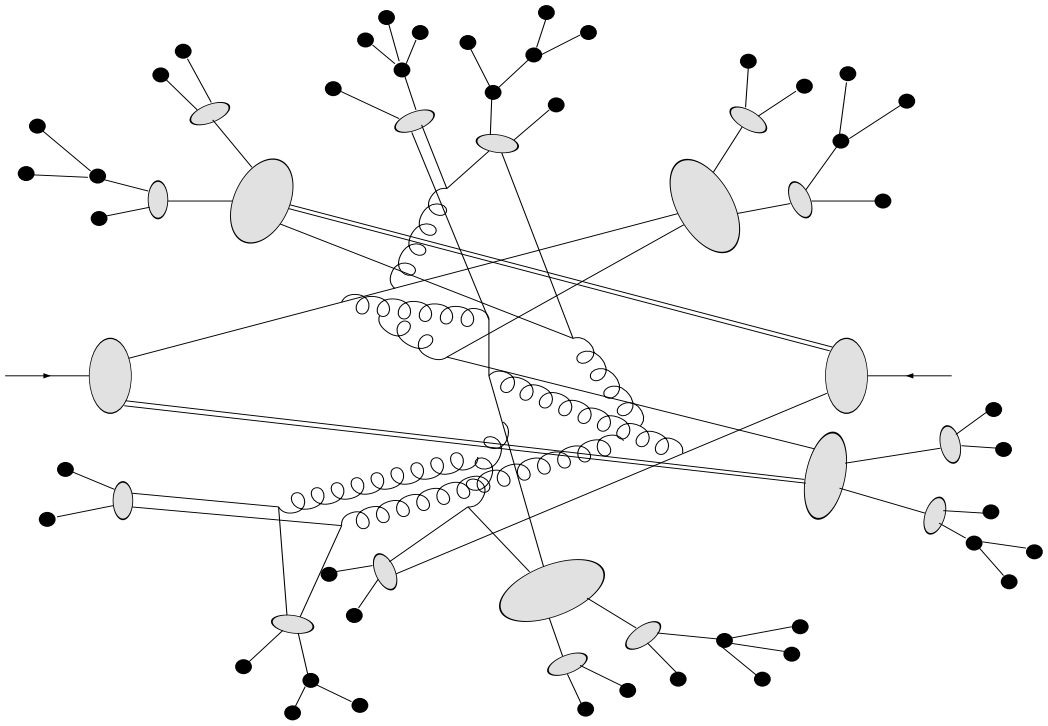

**Figure 1.6:** Cluster fragmentation. After preconfinement clusters are formed (blobs) which can further decay into new clusters of lower masses. Hadrons (black dots) are produced from these clusters.

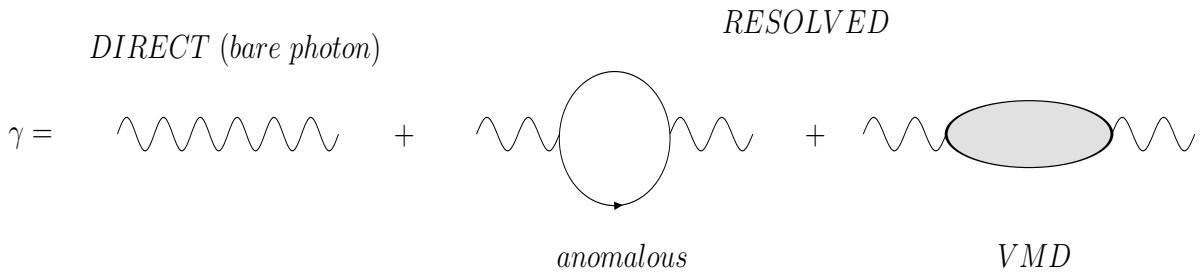


**Figure 1.7:** A photon can interact as bare point-like particle photon, as a $q\bar{q}$-fluctuation and as a vector meson with the proton.

$n(y, Q^2)$. The $ep$ cross section is written as:

$$d\sigma_{ep\to eX}(y, Q^2) = \int dy dQ^2 f_{\gamma/e}(y, Q^2) d\sigma_{\gamma p\to X}(y, Q^2) = \int dn(y, Q^2) d\sigma_{\gamma p\to X}(y, Q^2), \tag{1.20}$$

where $dn(y, Q^2) = f_{\gamma/e}(y, Q^2) dy dQ^2$. The flux of these quasi-real photons was studied independently by Weizsäcker [13] and Williams [14]. In the equivalent photon approximation the $Q^2$ dependence in the photon flux is kept:

$$f_{\gamma/e}(y, Q^2_{max}) dy = \frac{\alpha}{2\pi}\left[\frac{1+(1-y)^2}{y} ln\left(\frac{Q^2_{max}}{Q^2_{min}}\right) - 2m_e^2 y\left(\frac{1}{Q^2_{min}} - \frac{1}{Q^2_{max}}\right)\right] dy, \tag{1.21}$$

where $Q^2_{min} \approx \frac{m_e^2}{1-y}$ and $Q^2_{max} \approx (1-y)s$ unless experimental conditions reduce the range. Depending on the virtuality, $Q^2$, the photon can develop a complicated hadronic structure [15]. The life-time of an $e\gamma$-state is of the order $\sim 1/Q^2$. Within this life-time the photon can develop $q\bar{q}$-fluctuations. The life-time of these fluctuations are constant as a function of the characteristic $p_T$ of the interacting partons and are of the order $\sim 1/p_T^2$. Therefore, these fluctuations are important only if $Q^2 \ll p_T^2$. In this case, the photon can fluctuate into a $q\bar{q}$ pair or even more complicated states, where photon parton densities corresponding to vector mesons (VMD)[9] are introduced. These events have similar characteristics to hadron-hadron collisions.

The photon exchange can be then described by including all the possible photon states: as a bare, point-like particle photon, as a $q\bar{q}$-fluctuation and as a vector meson (figure 1.7).

The fraction of the photon energy entering in the hard scattering, $x_\gamma$, may be used to select events with resolved or point-like (direct) photons. At the LO parton level, $x_\gamma = 1$

[9] This is known as vector meson dominance. For quasi-real photons the vector meson state dominates over the perturbative $q\bar{q}$-fluctuations [16]. Other approaches are given in [17]

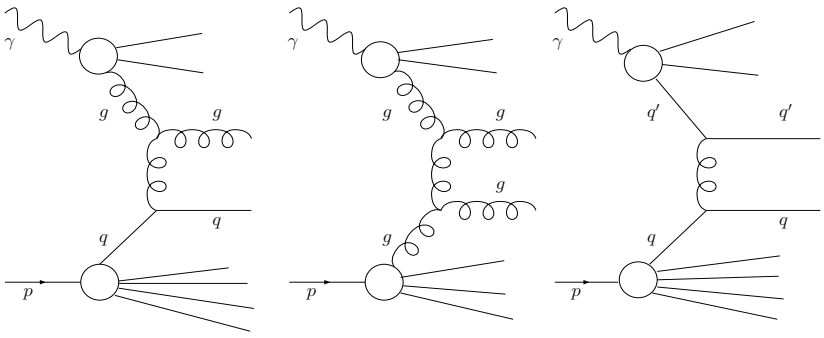


**Figure 1.8:** Resolved dijet LO QCD processes.

for direct processes whereas in the resolved case $x_\gamma < 1$. Experimentally, the variable $x_\gamma^{obs}$ is used. It is defined in terms of the two hardest jets, $Jet_1$ and $Jet_2$, and the hadronic final state (HFS) as:

$$x_\gamma^{obs} = \frac{\sum_{h \in Jet_1}(E - P_z) + \sum_{h \in Jet_2}(E - P_z)}{\sum_h (E - P_z)}, \tag{1.22}$$

where the sums in the nominator run over the hadrons in $Jet_1$ and $Jet_2$, while the sum in the denominator runs over all hadrons in the entire HFS. A derivation can be found in appendix A.

In the case of resolved photon events new diagrams, see figure 1.8, have to be added to the ones in figures 1.2 and 1.3.

The cross section is then assumed to factorize

$$\sigma = \sum_i \sum_j \int_0^1 \int_0^1 f_{i/p}(x_p, \mu^2) f_{j/\gamma}(x_\gamma, \mu_\gamma^2) \hat{\sigma}_{ij}\left(\hat{s}, \mu, \mu_\gamma, \mu_R\right) \; dx_p dx_\gamma, \tag{1.23}$$

where $\mu$ and $\mu_\gamma$ are the factorization scales of the proton and the photon, respectively, and $\mu_R$ is the renormalization scale.

## 1.5 Next to Leading Order

The leading order calculations (LO) can roughly describe the total cross section and give some qualitative results. However, there are large theoretical uncertainties from the renormalization and factorization scale dependencies since these scales are in principle arbitrary. They are usually set to a the typical scale in the process as $Q^2$ in DIS, heavy quark masses in heavy quark production or $P_T$ in high $p_T$ processes.

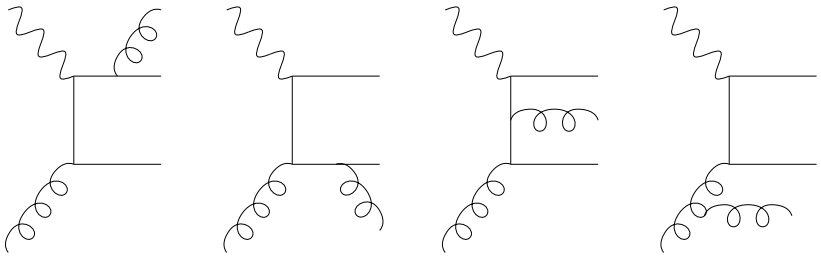

**Figure 1.9:** Diagrams considered at NLO for real emissions.

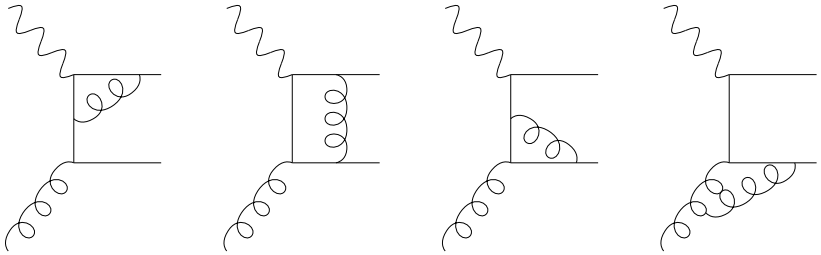

**Figure 1.10:** Virtual contributions considered at NLO (not all contributions are drawn).

By extending the calculations to NLO (figures 1.9 and 1.10) or even higher orders, we get more reliable predictions for the total and the differential cross sections. Therefore, from the theoretical point of view we can better test the kinematics of the process (section 1.6) and test the convergence of the perturbation series. Some other advantages are a better understanding of backgrounds and less sensitivity to the renormalization and factorization scales [18].

The improvement given by higher order corrections is especially important in the kinematic description of the events. If we consider dijet production at LO, the azimuthal angle between the jets is 180° whereas at NLO it can take values up to 120°. Adding higher order corrections can even improve further the kinematic description of partons and therefore of hadron and jet production.

## 1.6 Parton Evolution and Parton Showers

NLO calculations are not always available. On the other hand, parton evolution equations give the dominant contributions for parton radiation and therefore improve the event kinematics. There are several approaches to take into account the parton radiation depending on the assumptions and evolution variable used. The DGLAP, BFKL and the CCFM parton evolution equations are described in this section.

Equations 1.16 and 1.18 show how the quark, antiquark and gluon densities evolve as $ln\ Q^2$ changes through the splitting functions. The probability of parton $a$ to branch in a small interval of $\delta t$, where $t = ln\ q^2/\Lambda \rightarrow dt = dln(q^2) = dq^2/q^2$, $\delta t$, can be expressed with the splitting funcitons as:

$$\Gamma_a(t)\delta t = \int_{x_-(t)}^{x_+(t)} \sum_{b,c} \frac{\alpha_s}{2\pi}\ P_{a\rightarrow bc}(x)dx\delta t, \tag{1.24}$$

where $\Lambda$ is the $\Lambda_{QCD}$. The probability of no-branching in $\delta t$ is therefore: $1-\Gamma_a(t)\delta t$. Thus, the probability that a parton did not branch between $t$ and $t_0$ is the product of all probabilities of no-branching in all of the $\delta t$ intervals between these two virtualities. Adding all non branching contributions it can be expressed as $\Gamma_{a\rightarrow a}(t_\circ, t) = exp\left(-\int_{t_\circ}^{t} dt'\Gamma_a(t')\right)$. This can be used to calculate the probability for the parton $a$ at a virtuality $t_\circ$ of branching at virtuality $t$ :

$$\mathcal{P}_a(t) = \Gamma_a(t)\ exp\left(-\int_{t_\circ}^{t} dt'\Gamma_a(t')\right) \tag{1.25}$$

where $exp\left(-\int_{t_\circ}^{t} dt'\Gamma_a(t')\right)$ is the *Sudakov form factor*. This can be read as follows: the probability of branching for a parton by evolving from the virtuality $t_\circ$ to the virtuality $t$ is the probability of no-branching between $t_\circ$ and $t$ times the probability of branching at the virtuality $t$.

This notation is especially useful in Monte Carlo calculations since it describes the parton evolution in terms of increasing or decreasing virtuality. This allows to use it both for the incoming parton evolution towards the hard scattering and for the outgoing partons, i.e. *initial* and *final state parton showers*, respectively. Actually, Monte Carlo programs treat the initial state parton shower using the *"backward evolution"* [19, 20] approach in order to increase the calculation efficiency. Parton showers are not a NLO calculation but they resum all leading $log\ t$ terms to all orders as done in the parton evolution equations.

## DGLAP (Dokshitzer, Gribov, Lipatov, Altarelli and Parisi)

In the DGLAP equations, the parton evolution is described by a strong ordering in $t$, which leads to strong ordering in $k_t$, i.e the transverse momenta of the radiated partons. The strong ordering in $k_t$ means that the transverse momentum of the incoming parton into the hard scattering can be neglected compared to the transverse momentum in the hard scattering, figure 1.12 a). Therefore, in the collinear approach the partons incoming into the hard scattering are on-shell, so on-shell matrix elements are used.

The DGLAP equations describe most processes at HERA but since it resumes only $log\ t$ terms, it is expected to fail at small $x$ where contributions from $ln(1/x)$ also become large.

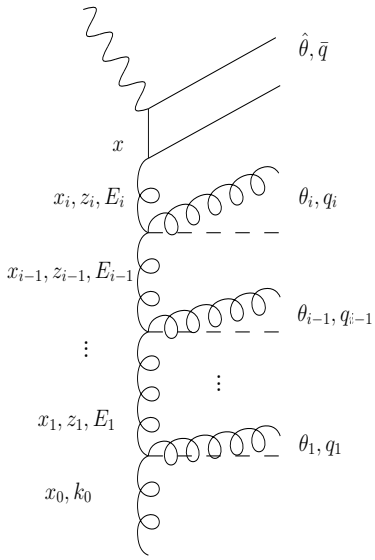


**Figure 1.11:** In the CCFM equation colour coherence or angular ordering is applied to the gluon ladder calculation.

## BFKL (Balitsky, Fadin, Kuraev, Lipatov)

In this approximation for large energies and small $x$, only leading contributions of $log\ 1/x$ are considered. We have therefore a strong ordering in the fractional momenta $x$: $x_0 \gg x_1 \gg ... \gg x_n$. Since there is no virtuality ordering, in this approximation $k_t$ can take any value and therefore off-shell ME are needed.

## CCFM (Catani, Ciafaloni, Fiorani and Marchesini)

The starting point of the CCFM equation [21,22,23,24] is the *angular ordering*, or *colour coherence*, in the emissions of a colour dipole[10] similar to a dipole in QED [25]. In the CCFM equation, angular ordering is applied to the gluon ladder (figure 1.11).

The transverse momentum of the $i$th emission is written as: $p_{t,i} = |q_i^\circ| \sin\theta_i$ and we define $z_i = E_i/E_{i-1}$. By energy conservation: $E_{i-1} = E_i + q_i^\circ = z_i E_{i-1} + q_i^\circ \Rightarrow q_i^\circ = (1-z_i)E_{i-1}$. Now, the transverse momentum of the $i$th emission is written as: $p_{t,i} = (1-z_i)E_{i-1}\sin\theta_i \simeq (1-z_i)E_{i-1}\theta_i$. With the definition of the *rescaled transverse momentum* $q_i = p_{t,i}/(1-z_i)$, the angle of the $i$th emission can be expressed as: $\theta_i = q_i/E_{i-1}$ and therefore $\theta_{i+1} = q_{i+1}/E_i$.

[10] Angular ordering is proven only at LO but not at higher orders.

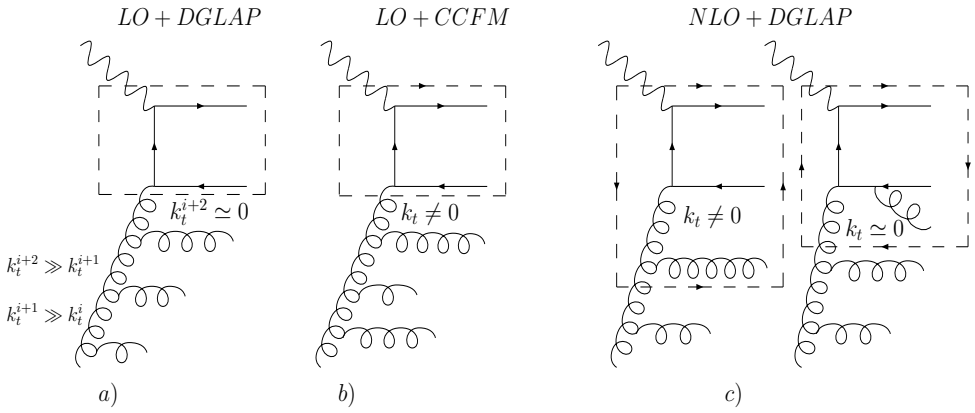


**Figure 1.12:** a) Leading order on-shell matrix element in $\alpha_s$ with DGLAP parton showers, b) leading order off-shell matrix element in $\alpha_s$ with CCFM parton showers and c) next to leading order on-shell matrix element in $\alpha_s$ with DGLAP parton showers c). Off-shell matrix elements with CCFM parton evolution include part of the NLO corrections.

Now, applying angular ordering, $\theta_{i+1} > \theta_i \Rightarrow q_{i+1}/E_i > q_i/E_{i-1}$ which can be rewritten as $q_{i+1} > z_i q_i$. For the full gluon ladder:

$$\bar{q} > z_n q_n,\ q_n > z_{n-1} q_{n-1}, \ldots,\ q_1 > Q_\circ \tag{1.26}$$

For moderate to small $x$ values $z_i \sim 1$ and angular ordering means $q_i > q_{i-1}$, which is the strong ordering in transverse momenta as in the DGLAP equations. For small $x$ instead, it leads to a non ordering in $k_t$ and since $k_t$ cannot be neglected, it is considered in all steps of the calculation.

This leads to the need of an *unintegrated parton density* (uPDF) $\mathcal{A}(x, k_t^2, \bar{q})$, where $k_t$ is treated explicitly. Since $k_t$ cannot be neglected, off-shell matrix elements are used. The factorization theorem, $k_t$-factorization theorem, can be found in [26, 27, 28].

An advantage of the CCFM parton evolution is that it does not neglect the $k_t$ of the incoming parton to the hard scattering. By considering $k_t$ on each splitting, it includes part of the NLO corrections (figure 1.12). Differences from the DGLAP evolution equations are seen in observables like forward jets

## 1.7 Jets

The partons from the hard scattering further radiate new partons which then hadronize. The energy flow is therefore blurred but at hadron level, the energy of a parton from the hard scattering is seen as a stream of energy collimated more or less in the direction of the original parton. This stream of energy is a *jet*. Jet algorithms are used to provide

a definition of jet and compare theory with experiments. Although there is no unique definition for jet finders they must fulfil the following conditions [29]:

- The definition of jet given by the algorithms should be infrared and collinear safe. When a parton radiates another parton with $|\vec{k}| \to 0$, infrared radiation, or at a very low angle $\theta \to 0$, collinear, the output of the jet algorithm should not change. As seen in section 1.3.2 the partonic cross sections were defined to be infrared safe and therefore jet algorithms have to ensure this property as well. Infrared and collinearsafety ensures a good correlation between parton and hadron levels.
- The sensitivity to the details of the hadronization should be small. Hadronization although not yet understood from first principles is expected only to smear the energy flow and therefore the jet algorithm should not be very sensitive to the details of the hadronization.
- They should be simple to use in experimental analyses and theoretical calculations.
- The jet algorithm should not be strongly affected by the remnants of the incoming particles and the underlying event.

In this analysis, the $k_\perp$ *jet-clustering algorithm* [29] is used. This algorithm fulfils the previous conditions and is invariant under boost along the beam direction. Because it is invariant under longitudinal boosts, it can be used in photoproduction in the laboratory frame because the hadronic centre-of-mass system is obtained by a longitudinal boost. The algorithm works as follows:

A - The algorithm starts with all the final state particles[11] $h_k$, here called objects, and computes the *resolution variables* $d_{kl}$ and $d_{kB}$ for all possible pair of objects $h_k$-$h_l$ and for objects and the beam $h_k$-B, respectively. The resolution variables used in this analysis are[12]:

$$d_{kB} = p_{t,k}^2 \tag{1.27}$$

$$d_{kl} = min\left(p_{t,k}^2, p_{t,l}^2\right)\left[(\eta_k - \eta_l)^2 + (\phi_k - \phi_l)^2\right]/R_\circ^2 \tag{1.28}$$

where $\eta_i = -ln\left(tan\frac{\theta_i}{2}\right)$, the pseudorapidity, and in this analysis $R_\circ = 1$.

B - The smallest $d_{kl}$ or $d_{kB}$ value is called $d_{min}$.

[11] Here particles means partons, hadrons or tracks and clusters depending on whether we run the jet algorithm at parton, hadron or detector level, respectively.

[12] The actual definition can change from implementation to implementation but at small angles $d_{kB}$ and $d_{kl}$ reduce to the transverse momentum of the $k$ object with respect to the beam direction and to the minimal relative transverse momentum of one object to the other, respectively.

C - If $d_{min}$ is the $d_{ij}$ variable, then the two objects $h_i$ and $h_j$ are combined into a new object $h_k$ which replaces $h_i$ and $h_j$. In this analysis the $k_\perp$ *jet-clustering algorithm* is used in the $p_t$-weighted recombination scheme[13]:

$$\begin{aligned} p_{t,k} &= p_{t,i} + p_{t,j} \\ \eta_k &= \frac{p_{t,i}\eta_i + p_{t,j}\eta_j}{p_{t,k}} \end{aligned} \tag{1.29}$$

$$\phi_k = \frac{p_{t,i}\phi_i + p_{t,j}\phi_j}{p_{t,k}} \tag{1.30}$$

D - If $d_{min}$ is the $d_{iB}$ variable, then the object $h_i$ is considered as a final jet and $h_i$ is removed from the objects considered by the algorithm.

E - The algorithm starts again with the remaining objects until no object is left.

[13]More examples on recombination schemes can be found in [29]. In this scheme masses are neglected and thus the resulting jets are massless.

# Chapter 2

# Multiple Parton Interactions

In this chapter we introduce multiple parton interactions (MPI) and the underlying event (UE) as a step further in the understanding of QCD. First we describe the general structure of Monte Carlo generators while giving a brief summary of most of the topics in chapter 1. Then, we present the multiple parton interactions and the experimental and theoretical motivations.

## 2.1 Monte Carlo event generators

In chapter 1 *ep* collisions were expressed as a convolution of parton density functions and a hard scattering, i.e. long and short distance contributions. This is supplemented with initial and final state parton showers. Finally, hadronization models for the fragmentation of partons into hadrons and their further decay into long lived hadrons are added (figure 2.1).

Monte Carlo generators (MC) take advantage of factorization theorems and make use of all our theoretical understanding. The different elements in the event simulation are:

- **Parton density functions** represent the probability to find a parton in the proton with a given momentum fraction $x$ at a scale $Q^2$, $F_2(x, Q^2)$, and in the unintegrated parton densities (uPDF) also with a given $k_t$, $\mathcal{A}(x, k_t^2, \bar{q})$. They cannot be derived from first principles and have to be obtained from data. This is done by adjusting a starting distribution such that after the parton evolution to large scales and a convolution with matrix elements (ME) data are described.
- The QCD **matrix elements** account for the hard scattering. In this analysis, the MC used have ME[1] only at leading order in $\alpha_s$. In the collinear case, on-shell ME are used while if the CCFM parton evolution is used off-shell ME are needed (section 1.6).

[1] An example on how to calculate these ME is given in appendix C.

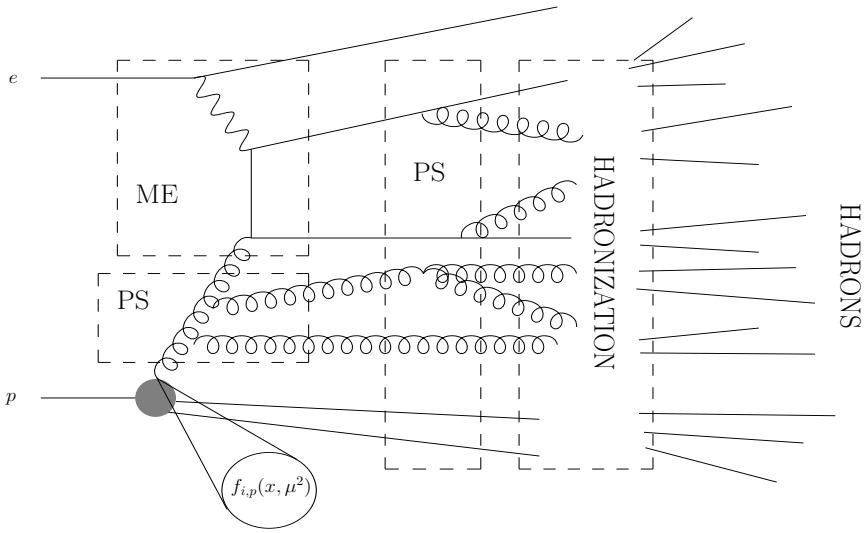


**Figure 2.1:** The *ep* cross section factorizes in a parton density functions and matrix elements (ME), i.e long and short distance contributions. This is supplemented with initial and final state parton showers and hadronization for the fragmentation of partons into colourless particles.

- **Parton showers**. By using the "backward evolution" partons are evolved from the (high) scales in the hard scattering to the (lower) scales in the proton while radiating partons, initial parton showers. Similarly, partons at high scales are evolved to lower scales in the final parton showers. Parton showers are important because higher order QCD radiation can be approximated by these means. Two parton evolution equations are available: DGLAP and CCFM.

- Finally, **hadronization** takes place and all partons fragment into colourless hadrons. Two hadronization models are used: the Lund string model, where hadronization is described in terms of strings between colour charges, and the cluster fragmentation model, where preconfinement leads to cluster which in turn decay into hadrons (section 1.3.4).

Three MC are used in this analysis: PYTHIA [30], with DGLAP parton evolution equations, on-shell LO matrix elements and the Lund string model for hadronization. HERWIG [9], with DGLAP parton evolution equation and LO matrix elements and the cluster fragmentation model. Finally, CASCADE [31, 32, 33], with the CCFM parton evolution equations an LO off-shell matrix elements with the Lund string model for hadronization.

## 2.2 Definition of the Underlying Event and Multiple Parton Interactions

In this section we define the underlying event (UE) and multiple parton interactions (MPI). Some examples of MPI search measurements from other experiments as well as at H1[2] are also given and finally, a theoretical model to account for MPI is presented.

Consider the case of two hadrons colliding, for example $p\bar{p}$. The ansatz used for resolved photons (i.e the photon parton density $f_{i/\gamma}$ is changed with the anti-proton parton density $f_{i/\bar{p}}$, cf. equation 1.23) can be used:

$$\sigma = \sum_i \sum_j \int_0^1 \int_0^1 f_{i/p}(x_p, \mu^2) f_{j/\bar{p}}(x_{\bar{p}}, \mu^2) \hat{\sigma}_{ij}\left(\hat{s}, \mu, \mu_R\right) \; dx_p dx_{\bar{p}}. \tag{2.1}$$

The parton density functions of the proton and anti-proton, the probability that one parton from each of the hadrons enters into the hard scattering, are convoluted with ME. This is supplemented with parton showers and remnants from the proton and the anti-proton are left, see figure 2.2 a).

This picture can be seen as a simplification since hadrons can be considered to be 'bunches' of partons. Therefore, when two hadrons collide, it is possible that more than a couple of partons interact, i.e. multiple partons interact (MPI). These are partons from the remnants as shown in figure 2.2 b). MPI from re-scatters off the remnants can also happen.

The underlying event (UE) is defined as everything except the lowest order process. The UE is therefore the initial and final state radiation and the remnant-remnant interactions as well as possible re-scatters off the remnants.

In high energy collisions, Lorentz contracted hadrons may contain many partons. This is because at high energies, low $x$ values become accessible and the density of partons grows with decreasing $x$. In this dense system of partons, many of them could interact.

There are many challenges in the theoretical description of such events. Which parton density functions to use? How dependent/independent are the different scatterings? Are they colour connected? Is it really necessary to take them into account?

### 2.2.1 Previous Measurements at Hadron Colliders

With four jet production, double parton scattering can be studied (figure 2.3). Jet correlations are different in case of MPI b1) and c1) or parton radiation b2) and c2).

In events where MPI are present, each parton scattering balances pairwise in momentum, i.e. ${P_\perp}^a + {P_\perp}^b \simeq 0$ and ${P_\perp}^c + {P_\perp}^d \simeq 0$ but there is no correlation between each scattering.

[2]The H1 experiment is presented in chapter 3.

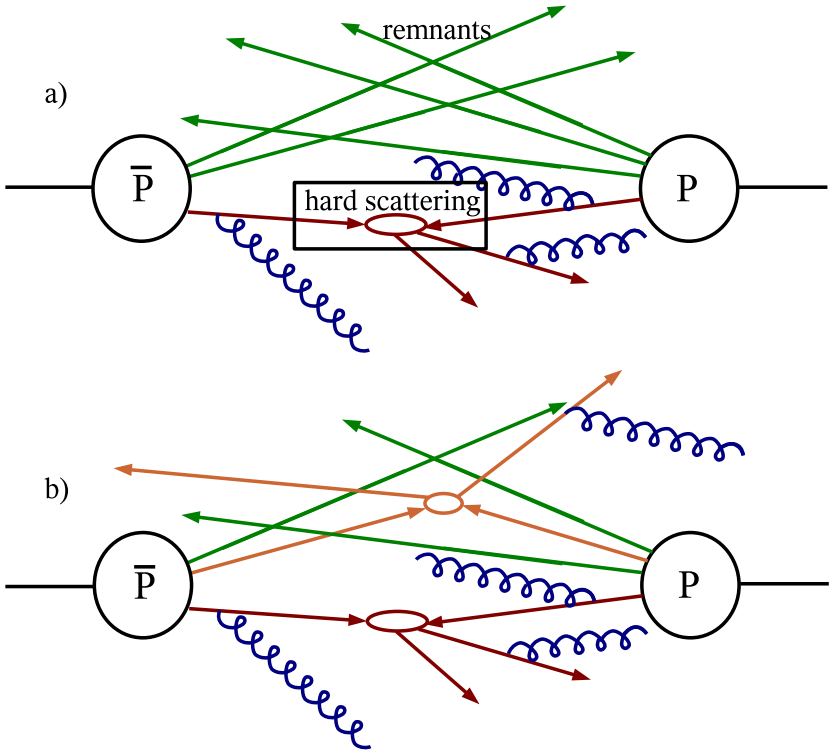


**Figure 2.2:** Parton density functions of the proton and anti-proton are convoluted with ME and supplemented with parton showers. a) Hadron remnants are left and b) the partons from the remnants can interact (MPI).

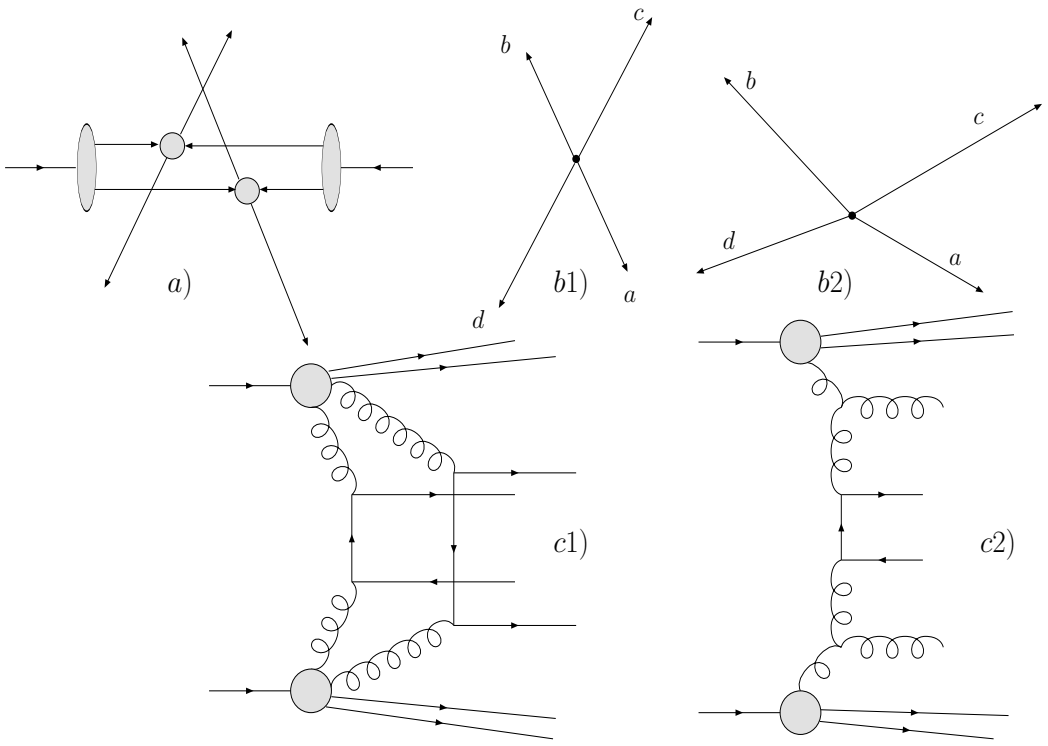


**Figure 2.3:** Four jet production from MPI in hadron collisions a). The jet correlation is different in case of MPI, where each collision balances in momentum b1) and c1), or in case of parton radiation b2) and c2), where the total momentum balances but not pairwise.

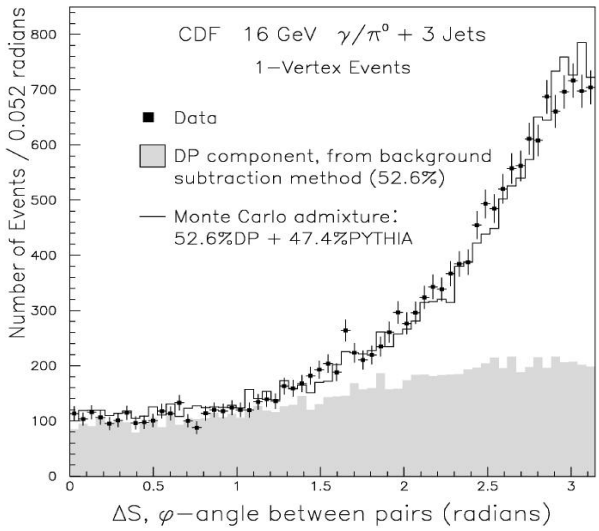


**Figure 2.4:** 3 jet + prompt photon production [34] ($P_T^{Jets} > 5$ GeV and $P_T^{\gamma} > 16$ GeV) as a function of $\Delta S = \phi_i - \phi_j$. Events with MPI give a rather flat contribution (shaded) while events without MPI contribute at high values of $\Delta S$.

If only one hard scattering takes place together with parton radiation, only the total momentum balances, i.e. $P_\perp{}^a + P_\perp{}^b > 0$ and $P_\perp{}^c + P_\perp{}^d > 0$ but $P_\perp{}^a + P_\perp{}^b + P_\perp{}^c + P_\perp{}^d \simeq 0$.

The **CDF collaboration** studied 3 jet + prompt photon[3] production ($P_T^{Jets} > 5$ GeV and $P_T^{\gamma} > 16$ GeV) [34]. They analyzed the $\Delta S$ distribution, where $\Delta S = \phi_i - \phi_j \; \forall$ pair of jets or jet and photon $i \wedge j$. This is shown in figure 2.4. In events with MPI the $\Delta S$ distribution is rather flat (shaded). This is because most of the $\Delta S$ contributions from these events are from uncorrelated pairs, i.e. double parton scattering. However, if the jets are produced from hard parton radiation, the angular separation is always correlated and usually large.

Figure 2.4 shows a large contribution of double parton scattering as given by the PYTHIA MC generator (shaded). This contribution is especially large at low $\Delta S$ values. This distribution is considered one of the strongest evidences of MPI.

Another example is given in figure 2.5. The **UA5 collaboration** measured [35] the charged particle multiplicity and forward-backward correlations in the detector by considering two bins in pseudorapidity. The first bin, between $\Delta\eta/2$ and $\Delta\eta/2+1$ (forward) and the second, between $-\Delta\eta/2$ and $-\Delta\eta/2+1$ (backward). The variable $b(\Delta\eta) = \frac{\langle n_F n_B\rangle - \langle n_F\rangle^2}{\langle n_F^2\rangle - \langle n_B^2\rangle}$ was defined, where two one-unit-wide bins are separated by a central gap $\Delta\eta$.

Figure 2.5 shows that besides normalization, predictions without MPI (left) do not de-

[3]Prompt photon are produced in the hard scattering process and experimentally are characterized by an isolated photon with high transverse momenta.

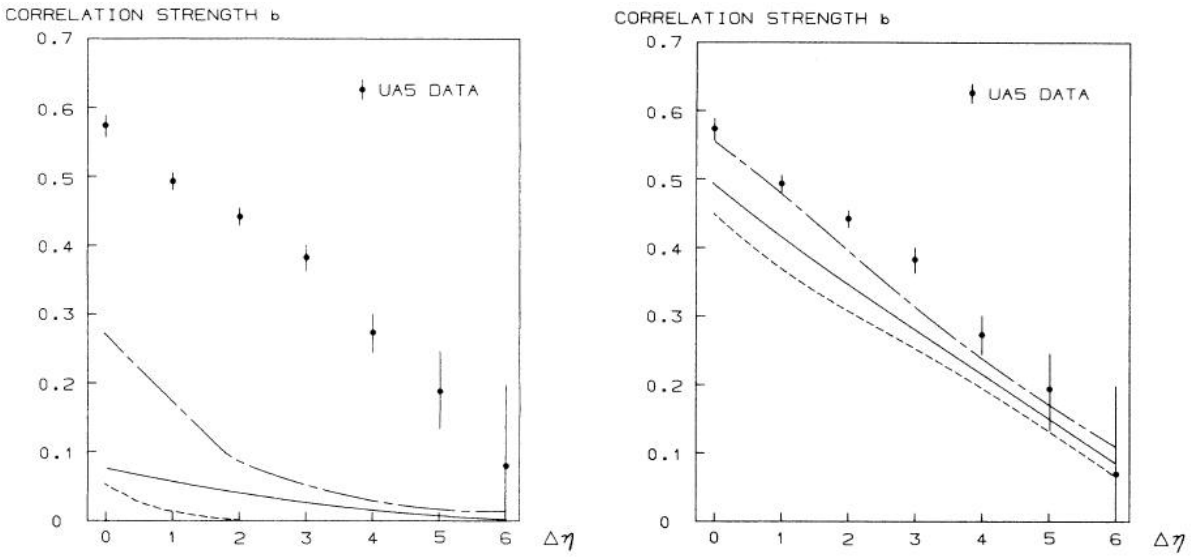


**Figure 2.5:** Charge particle multiplicity forward-backward correlations at the UA5 collaboration [35]. The correlations $b(\Delta\eta) = \frac{\langle n_F n_B\rangle - \langle n_F\rangle^2}{\langle n_F^2\rangle - \langle n_B^2\rangle}$ cannot be described by models without MPI models (left) while models which include MPI (right) can.

scribe the shape of the distribution. If in an event only a single scattering is present, particles are mainly produced in a given $\eta$ region while out of this region only a few particles are produced. If MPI are present (right), besides the particles produced by the hardest interaction, more particles are produced in other $\eta$ regions of the detector. Consequently, forward and backward hemispheres are simultaneously influenced in presence of MPI, i.e. correlations in the different regions of the detector are higher.

### 2.2.2 Previous Measurements at HERA

The H1 and ZEUS collaborations studied MPI at HERA. At **H1**, photoproduction events with two jets were studied [36]. The average transverse energy density, $\langle E_T\rangle/(\Delta\eta\Delta\phi)$, in the $|\eta^*| < 1$ region outside the two jets was measured as a function of $x_\gamma^{obs}$ (equation 1.22). The results are shown in figure 2.6. Again, models not including MPI (PYTHIA) cannot describe the measurement while models which include MPI can, for example PYTHIA mia.

Since most of the energy from the hardest interaction is inside the two leading jets, the transverse energy outside the two jets is expected to be sensitive to MPI. In this measurement the additional activity is well described when adding MPI. Towards lower $x_\gamma^{obs}$ values, the transverse energy grows since the resolved component of photon becomes more important.

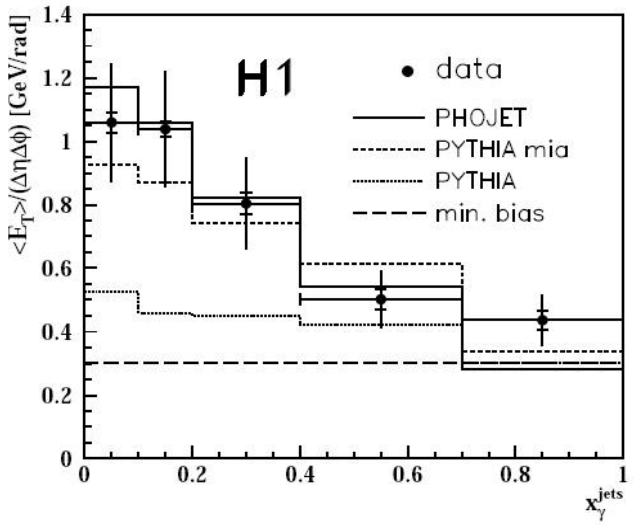


**Figure 2.6:** Average transverse energy density, $\langle E_T \rangle / (\Delta\eta\Delta\phi)$, in the $\eta^* < 1$ region outside the two jets as a function of $x_\gamma^{obs}$ [36].

The **ZEUS** collaboration studied MPI with four jet photoproduction events [37] ($E_T^{Jet_{1,2}} >$ 6 GeV and $E_T^{Jet_{3,4}} > 5$ GeV). Figure 2.7 a) shows the event distribution as a function of $x_\gamma^{obs}$. The direct contribution (shaded) dominates in the high $x_\gamma^{obs}$ region. Including contributions from resolved photons, PYTHIA without MPI (dashed line) and PYTHIA with MPI (solid line) describe data at high $x_\gamma^{obs}$ values ($x_\gamma^{obs} > 0.6$). At low $x_\gamma^{obs}$ values the data are described when MPI are included. The contribution from MPI increases with decreasing $x_\gamma^{obs}$.

Figure 2.7 b) shows the event distribution as a function of $x_\gamma^{obs}$ where the four jets invariant mass is larger than 50 GeV ($m_{4J} > 50$ GeV). The jet invariant mass selection is expected to remove non-perturbative contributions in the events. This is in agreement with the distribution and the description given by the PYTHIA MC. It describes the data when including the contributions from resolved photons regardless of whether one uses MPI models or not.

In conclusion, experimentally MPI have been observed at hadron colliders and are needed to describe the data. At HERA, MPI can be studied as a function of the energy fraction of the photon entering into the hard scattering $x_\gamma^{obs}$. At high $x_\gamma^{obs}$ values the photon is a point-like particle while at low $x_\gamma^{obs}$ values it shows a hadronic structure and MPI are possible. At low $x_\gamma^{obs}$ values, models including MPI describe the data while if MPI are not included the data cannot be described.

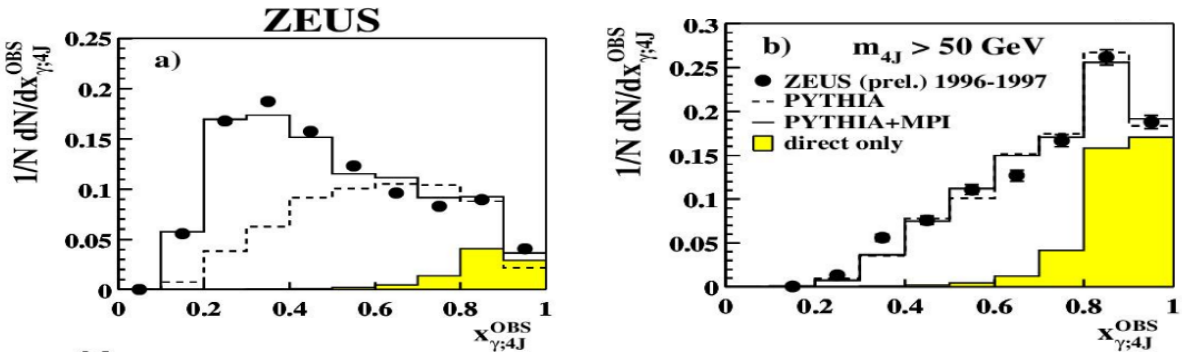


**Figure 2.7:** a) Four jet event distribution [37] as a function of $x_\gamma^{obs}$ and b) four jet with invariant mass $m_{4J} > 50$ GeV event distribution. The invariant mass selection removes non-perturbative contributions in the events.

## 2.3 Multiparton Interaction Models

The rapid rise of the parton density at low $x$ (figure 1.4) suggests a proton viewed as a 'bunch' of partons. This leads to the possibility that multiple parton collision from two incoming hadrons can occur. Actually, with increasing centre-of-mass energies the momentum fraction $x$ probed are smaller. Since the parton densities at small $x$ are larger, the probability of MPI increases. Eventually, the number of partons inside the hadron can grow enough so that partons may recombine, i.e. the number of partons *saturate*[4].

Several models for MPI exist. MC generators which include MPI models are HERWIG/JIMMY, and PYTHIA. The MC PYTHIA offers a rather simple model in $ep$ collisions, which will be presented here. The HERWIG/JIMMY MPI model will be shortly described.

### 2.3.1 The PYTHIA Multiparton Interaction Model

The starting point in PYTHIA is the observation that the perturbative QCD (pQCD) 2→2 cross section[5] diverges like $d\sigma/dp_\perp^2 \approx 1/p_\perp^4$ for $p_\perp \to 0$, where $p_\perp$ is the relative transverse momentum between the two outgoing partons. This leads to an increasing integrated cross section $\sigma_{int} = \int_{p_{\perp,min}}^{\sqrt{s}/2} \frac{d\sigma}{dp_\perp} dp_\perp$ with decreasing $p_{\perp,min}$ values. Thus, the integrated cross section can become larger than the total inelastic non-diffractive cross section, $\sigma_{nd}$. In figure 2.8 it is shown for $pp$ collisions at $\sqrt{s} = 14$ TeV and $p\bar{p}$ collisions at 1.8 TeV [38] (left), and for $ep$ collisions at $\sqrt{s} = 318$ GeV (right).

This is not a contradiction if $\sigma_{int}$ is considered to be not the hadron-hadron cross section

[4]Saturation effects have no yet been unambiguously observed.

[5]The cross section for: $qq' \to qq'$, $q\bar{q} \to q'\bar{q}'$, $q\bar{q} \to gg$, $qg \to qg$, $gg \to gg$ and $gg \to q\bar{q}$. These processes are dominated by the t-channel which has this characteristic.

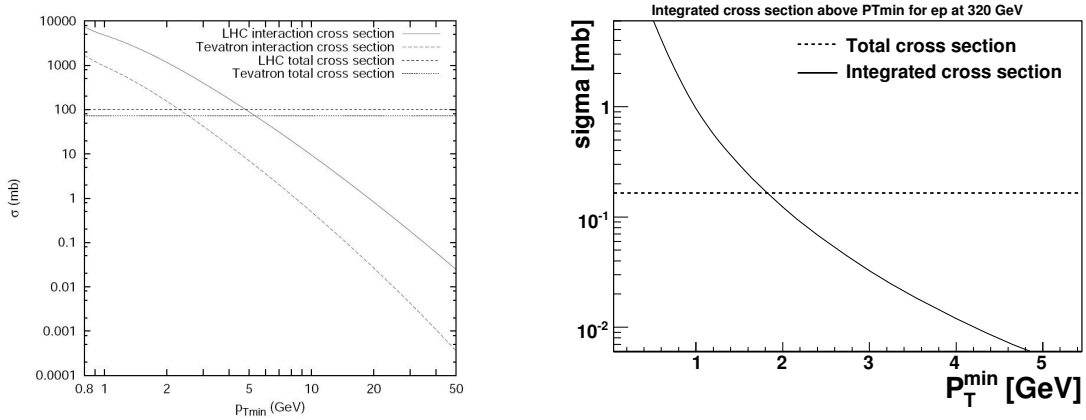


**Figure 2.8:** Integrated cross section at 14 TeV for $pp$ collisions and at 1.8 TeV for $p\bar{p}$ collisions (left) [38] and the integrated cross section for $ep$ collisions at 318 GeV as a function of $p_{\perp}^{min}$ (right) (done numerically using the CTEQ 5L and the CTEQ 6L parton density functions, respectively). The integrated cross section will eventually exceed the total inelastic non-diffractive cross section.

but the parton-parton one, i.e. $n$-multiple parton interactions contribute $n$ times to the integrated cross section but only once to the total cross section. The partonic cross section can be related to the inelastic one by

$$\langle n \rangle (p_{\perp,min}) = \frac{\sigma_{int}}{\sigma_{nd}}, \tag{2.2}$$

which leads to an average number of partonic collisions per event larger than one and where now $p_{\perp,min}$ is a parameter of the model[6]. A direct consequence of introducing MPI is that we allow lower $p_{\perp}$ values in the model. In the limit $p_{\perp} \to 0$, the number of MPI becomes infinite which is not possible[7]because this would mean that the scattered energy becomes infinite.

Energy-momentum conservation has to be introduced so that after one interaction the energy left for further interactions decreases and eventually no energy is left. Therefore, the introduction of multi-parton correlated parton distributions could tame the divergency but these are not available since they are not part of the standard pQCD.

The reason for an effective cut-off is the following. In scattering theory, the incoming and outgoing particles are considered as free particles. On the contrary, quarks and gluons are

[6]In the $k_{\perp}$ factorization formalism the soft emissions do not contribute to the total cross section while contributing to the event properties since off shell matrix elements are used. Therefore, in this formalism no cut-off would be needed.

[7]The parameter has to be tuned to data and at current collider energies is of the order of $1-2$ GeV. At larger energies this value is expected to be higher.

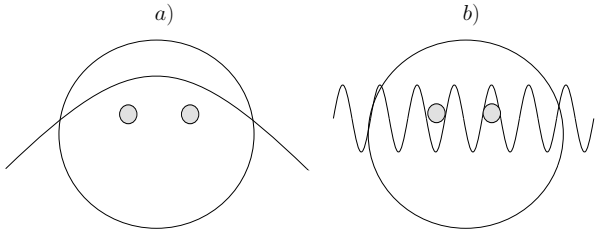


**Figure 2.9:** Depending on the wave length of the probing particle ($\sim 1/p_\perp$) two colour charges can be resolved.

not free before and after scattering due to confinement, i.e. they are not asymptotically free, and are in colour neutral states. At low $p_\perp$, the time of interaction is too large to neglect confinement and therefore pQCD cannot be applied. If a coloured particle with a given $p_\perp$ probes the proton, the question of whether it can resolve the colour charges or not depends on the wave length of the probing particle ($\sim 1/p_\perp$), figure 2.9. Therefore, at low $p_\perp$ the probing particle cannot resolve the structure of the proton, i.e. it cannot 'see' the colour charges, and therefore will not interact, dampening the divergency at low $p_\perp$ values. Therefore, an effective cut-off is motivated and the model is consistent.

The model considers each parton scattering to be independent. This leads to a poissonian[8] probability distribution of the MPI with mean $\langle n \rangle$ (from equation 2.2). Since the incoming hadrons collide at very high energies, they can be seen as a 'pancake' with its components 'frozen' due to Lorentz contraction, which supports the hypothesis of independent MPI.

For Monte Carlo implementation it is useful to define:

$$f(x_\perp) = \frac{1}{\sigma_{nd}(s)} \frac{d\sigma}{dx_\perp}, \tag{2.3}$$

the probability to have a parton-parton interaction at $x_\perp = 2p_\perp/E_{cm}$ given that the two hadrons undergo a non-diffractive inelastic collision. Using equation 2.3, the probability that the hardest interaction, i.e. with higher $x_\perp$, is at a given $x_{\perp,1}$ value can be written as

$$\mathcal{P}(x_{\perp,1}) = f(x_{\perp,1})\ exp\left(-\int_{x_{\perp,1}}^{1}, f(x'_\perp)dx'_\perp\right), \tag{2.4}$$

i.e. the probability that there was no scattering with $x_\perp$ larger than $x_{\perp,1}$ multiplied by the probability that a scattering at $x_{\perp,1}$ occurs. Note that this expression is similar to

[8]The Poisson distribution has only one paremeter, the average value $\mu = \langle n \rangle$. Here, some corrections are needed from $n_{int} = 0$ number of interactions since for an inelastic scattering at least one interaction is supposed to happen.

Sudakov form factor in equation 1.25. Therefore, within the formalism of the model, an ordering in $x_\perp$ arises naturally: $1 > x_{\perp,1} > x_{\perp,2} > ... > x_{\perp,i}$.

For the next $i$th scattering it can be shown that its probability is given by:

$$\mathcal{P}(x_{\perp,i}) = f(x_{\perp,i}) \frac{1}{(i-1)!} \left( \int_{x_{\perp,i}}^{1} f(x'_\perp) dx'_\perp \right)^{i-1} exp\left( - \int_{x_{\perp,i}}^{1} f(x'_\perp) dx'_\perp \right). \quad (2.5)$$

The parton density functions are used for the hard scattering[9]. Since no multi-parton correlated parton distributions are available, after the hardest interaction the parton densities are modified to account for the available phase space and the parton flavour kicked out in the hard scattering. To account for the energy used after $i-1$ scatterings, for the $i$th scattered parton the $x'_i$ momentum fraction is used instead of $x_i$: $x'_i = \frac{x_i}{1-\sum_{j=1}^{i-1} x_j}$

In an event with MPI the possible colour connections between partons become complex and cannot be easily predicted by pQCD. The colour structure can influence the structure of the events. In the simple model considered here, only three possibilities are considered for the subsequent interactions:

- $gg \rightarrow gg$ with the two gluons in a colour-singlet state. Here, the two gluons are colour connected and colour decoupled from the rest of the event, figure 2.10 a).
- $gg \rightarrow gg$ with colour reconnections to one of the 'already' existing strings such that the new string length is as short as possible, figure 2.10 b).
- $gg \rightarrow q\bar{q}$ in colour-singlet state and therefore the colour connection is stretched between the outgoing $q\bar{q}$.

Colour string configurations like the one shown in figure 2.10 a) maximize the total string length compared to b) and therefore more particles are produced.

In the model used in this analysis, initial and final state radiation is provided only for the hardest parton interaction. For the subsequent interactions no parton shower is included.

In principle, the model described here referred to hadron-hadron collisions. However, in lepton-proton collisions, the exchanged photon can develop a hadronic structure when its virtuality is low, i.e. in the photoproduction regime ($Q^2 < 1$ GeV$^2$) (section 1.4). Similar to the proton, parton density functions can be extracted for the photon. The different photon states, point-like, $q\bar{q}$-fluctuation and vector meson, are explicitly treated in the SaS [16] photon parton density functions and will be used in this analysis.

[9] The hardest collision is also the first one to be computed since within an event the MPI are considered in decreasing order in $x_\perp$ similarly to the parton showers.

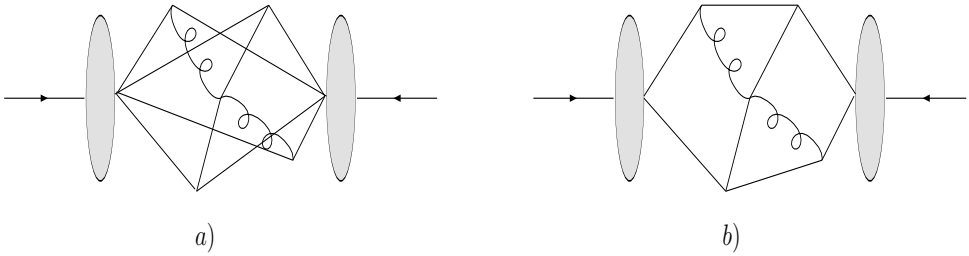


**Figure 2.10:** Possible colour connexions in MPI events for the simple model: a) colour reconnections to one of the 'already' existing strings and b) colour-singlet states colour decoupled from the rest of the event.

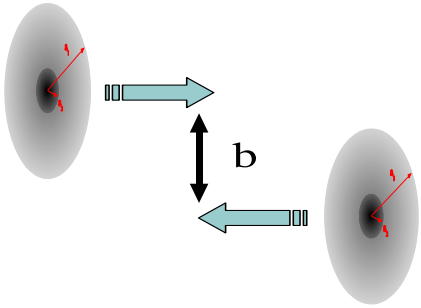


**Figure 2.11:** Hadrons are extended objects and an impact parameter, $b$, can be introduced. The matter distribution inside the hadron can be represented as Gaussian or double Gaussian.

Further developments of this model are available only for hadron-hadron collisions[10]. Here, the hadron is treated as an extended object and therefore an impact parameter, $b$, is introduced, i.e. a distance between the photon hadronic system centre and the proton centre (figure 2.11). A small impact parameter means central collision and high overlap of the hadron when colliding. Consequently, in this case, the MPI are more likely to occur. For each impact parameter value, the number of interactions is assumed to be distributed according to a Poissonian.

The matter distribution inside the hadron can be represented as Gaussian or double Gaussian, with a 'warm core' in the centre.

Finally, parton showers can be introduced in all remnant-remnant scatterings. In this case, the available phase space has to be shared between the parton showers and the remnant-remnant scatterings. In this extension of the model, the remnant-remnant scatterings and the initial parton showers are interleaved and start a phase space competition.

[10]The difference between hadron-hadron and lepton-hadron collisions is that in lepton-hadron collisions the photon-proton centre-of-mass energy in the photoproduction regime is $W^2 \simeq y\ s$. This means that in each collision the centre-of-mass energy is different which is not the case of hadron-hadron collisions.

### 2.3.2 The JIMMY Multiparton Interaction Model

Details about the JIMMY model for MPI can be found in [39,40]. Similarly to PYTHIA, JIMMY assumes that multiple parton interaction occur independently and therefore they obey Poisson statistics. It assumes a matter distribution and thus, a matter overlap function $A(b)$ which depends on the impact parameter. MPI are only considered when the photon can be considered as a non-perturbative bound state, i.e. the photon is described by vector meson dominance (section 1.4).

If not taken into account the total energy participating in the MPI can be larger than the initial energy from the incoming particles. Imposing energy-momentum conservation means that in practice the MPI are not really independent. In JIMMY, the energy-momentum conservation is imposed with a $\Theta$-function in the cross section so that the total energy is less than the initial energy. If after $n$ scatters there is no more available energy, further MPI are vetoed. This imposes a minimal correlation between the MPI.

In an event with MPI the colour flow of partons from the hard scattering are treated independently from the subsequent scatters. For the subsequent scatters no initial state radiation is generated and only final state radiation is simulated.

## 2.4 Multiple Parton Interaction Analysis Observables

When studying MPI different observables can be used. High $P_t^{Jets}$ can be used, for example 3 jet + prompt photon or 4 jet events used at the CDF and ZEUS collaborations, respectively (sections 2.2.1 and 2.2.2). Since the sensitivity of the jet algorithms to the details of hadronization is small, this observable is rather insensitive to hadronization. The higher the $P_t^{Jet}$ of a jet is, the less sensitive to hadronization effects are. On the other hand, this observable is directly sensitive to (semi-)hard MPI.

Another observable is particle production, for example the $b(\Delta\eta)$ particle correlations and the event distribution as a function of $x_\gamma^{obs}$ at the UA5 and H1 collaborations, respectively. In events where MPI are present, more particles are produced and the production mechanisms can be tested by measuring particle production in different regions of the event and/or their correlations. The UA5 collaboration used the $b(\Delta\eta)$ distribution to look at correlations between different $\eta$ regions and the H1 collaboration measured particle production outside the leading jet, i.e. the hard interaction. This observable is not only sensitive to MPI but also to the hadronization details (1.3.4). It is sensitive both to the different colour flow configurations, i.e. string lengths and preconfinement, and to the string and cluster dynamics in the Lund and the cluster fragmentation models, respectively.

In this analysis two observables are used: charged particles and low $P_T$ jets, minijets, multiplicity. The dijet event, particle and minijet selection are given in sections 4.1 and 4.2.

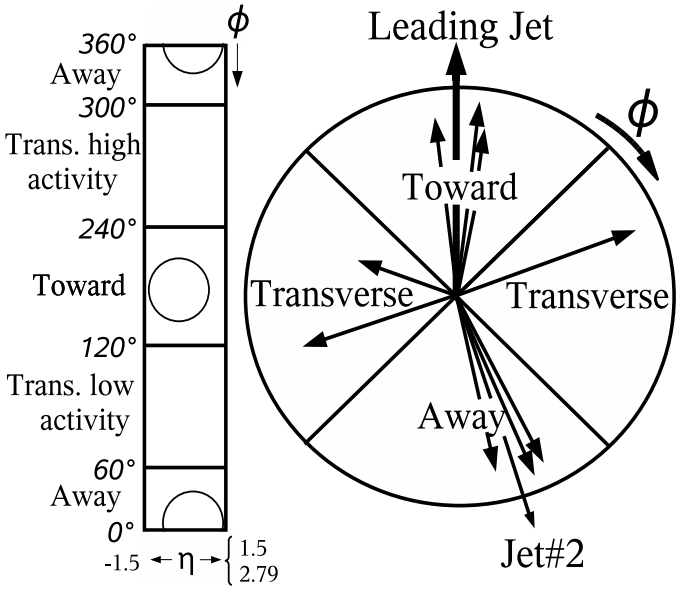


**Figure 2.12:** The leading jet defines the *toward region* in the transverse plane. The azimuthal angle of the observables, i.e charged particles and minijets, are measured from the leading jet, which is set to be at $\phi = 180°$.

The leading jet defines the *toward region* in the transverse plane. The azimuthal angle of the observables, i.e charged particles and minijets, from the leading jet, which is set to be at $\phi = 180°$ (figure 2.12), defines different regions in the transverse plane: the *away* and *transverse* regions. The second leading jet is usually located in the away region, although not necessarily especially when minijets are present (momentum conservation).

In each hemisphere of the transverse region the eventwise scalar sum of the transverse momenta is measured:

$$P_T^{sum} = \sum_{i}^{\#obs} P_T^i, \tag{2.6}$$

where *#obs* is the number of observables, tracks or minijets, in a given event. Comparing the $P_T^{sum}$ of each hemisphere the transverse high and low activity regions are defined: the hemisphere with high $P_T^{sum}$ is the high activity region while the other hemisphere is the low activity region.

Summarizing, the definition of the regions in the transverse plane is:

- Toward: $120° < \Delta\phi < 240°$
- Away: $300° < \Delta\phi < 60°$

- Transverse low: $60^\circ < \Delta\phi < 120^\circ$
- Transverse high: $240^\circ < \Delta\phi < 300^\circ$

If in an event the transverse high activity region is in the region defined for low activity region and vice versa, the event is rotated to match the above definition.

Already in previous measurements at HERA, the variable $x_\gamma^{obs}$ was used in order to select events with resolved photons, sections 1.4 and 2.2.2. In this analysis, events with enhanced resolved photon contribution are selected when $x_\gamma^{obs} < 0.7$. The point-like enhanced contribution events are events where $x_\gamma^{obs} > 0.7$.

The definition of the regions in the transverse plane separates the hard scattering, toward and away regions, from the event. The purpose of defining the high and low activity transverse regions is to separate possible higher order contributions from the underlying event. If higher order contributions go to a given hemisphere of the transverse region it adds extra $p_T$ to that hemisphere. Therefore, the transverse high activity region is supposed to receive most of the higher order effects. This can be shown by comparing the average number of charged particles in the transverse regions with and without parton showers using a MC generator.

This is shown in figure 2.13, where the contributions from initial (Initial SR) and initial and final state radiation (Parton showers) are compared to the MC without any parton shower (No PS). The parameters in PYTHIA can be found in appendix D. The figure shows the average particle multiplicity in the high and low activity regions (upper and lower plots, respectively) as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$. In the high activity region this contribution is in average about 0.66...0.78 and 0.3...0.13 charged particles for low and high $x_\gamma^{obs}$ values, respectively. In the low activity region the average amount is less, about 0.35...0.46 and 0.13 charged particles.

The **average track multiplicity** is measured **as a function of** the azimuthal angle with respect the leading jet, which is set to be at $\phi = 180^\circ$, $\mathbf{\Delta}\phi^{Jet_1,\ h\pm}$, in the two $\mathbf{x}_\gamma^{obs}$ regions above defined.

Both the **tracks and minijets average multiplicity** are measured **as a function of** the transverse momentum of the leading jet, $\mathbf{P}_T^{Jet_1}$ in the two $\mathbf{x}_\gamma^{obs}$ regions above defined.

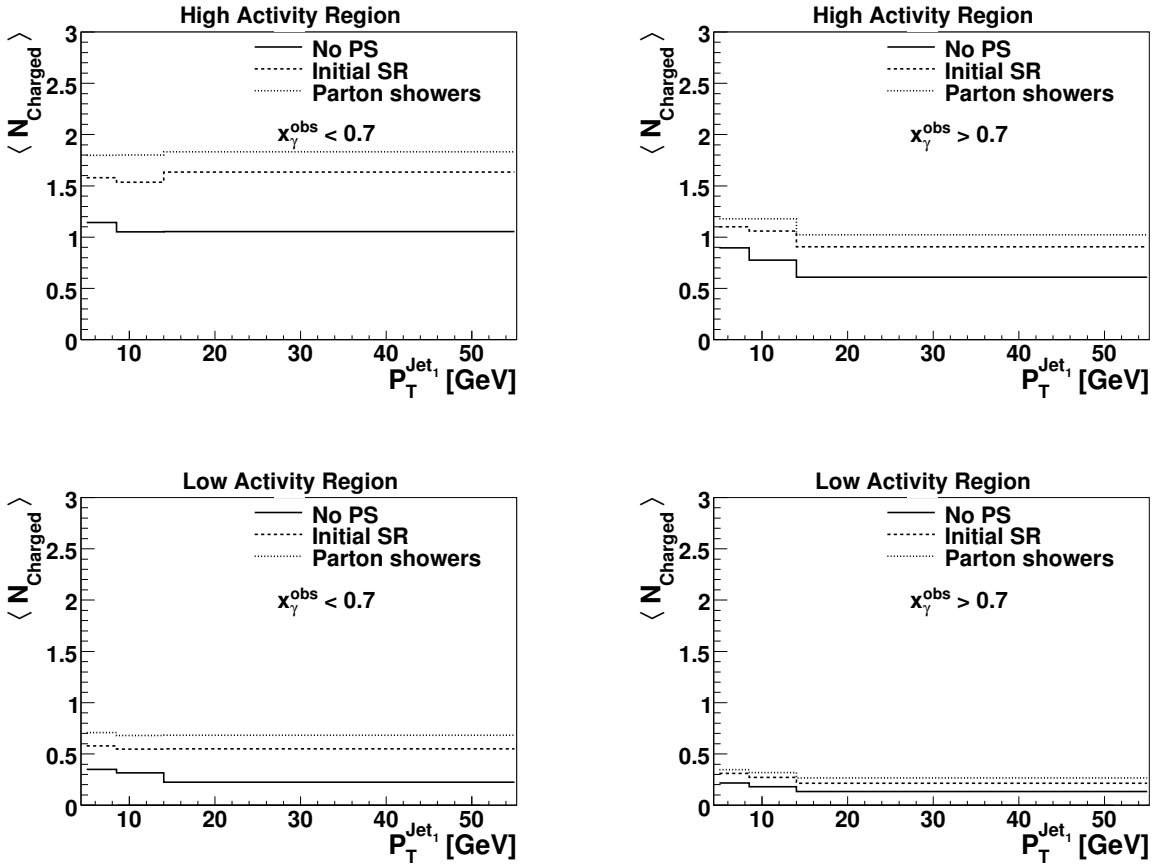


**Figure 2.13:** Contributions from initial (Initial SR) and initial and final state radiation (Parton showers) are compared to the MC without any parton shower (No PS). The contribution of parton showers is higher in the high activity regions than in the low activity regions for low and high $x_\gamma^{obs}$ values.

# Chapter 3

# The H1 Experiment at HERA

The data used in this analysis were taken with the H1 detector, one of the experiments at the HERA[1] storage ring at the DESY[2] laboratory, in Hamburg (Germany). This chapter introduces HERA and gives the necessary description of the main components of the H1 detector needed in this analysis.

## 3.1 HERA

HERA was a lepton-proton accelerator. It was the first and the only one of this type in the world to date . Its construction started in 1984 and on 19th of October 1991 the first collisions took place. HERA operated until the end of June 2007. The ring had a circumference of 6336m. Electrons[3] were accelerated up to 27.6 GeV and protons up to 820 GeV before 1998 and later up to 920 GeV.

The two beams collided at two points: at the H1 experiment, in the Hall north, and at the ZEUS experiment, in the Hall south, see figure 3.1. Besides of these two experiments there were two fixed target experiments: HERMES, which used longitudinally polarised electrons, and HERA-B (until 2003) which used protons. HERMES was designed to study the spin structure of nucleons and HERA-B to measure CP violations in B meson decays.

The experiments were designed to perform precision measurements, among them the determination of parton densities and very precise test of the standard model. In the last months of operation the proton beam was accelerated to 460 GeV and 575 GeV, the so called low and medium energy runs. The goal in these last months was to perform a direct measurement of $F_L$. In the analysis described here, the data were taken in two years: 1999 and 2000, where the centre-of-mass energy was $\sqrt{s}$ ≈320 GeV.

[1] From the German abbreviation: **H**adron-**E**lektron-**R**ing-**A**nlage.
[2] From the German abbreviation: **D**eutsches **E**lektronen **Sy**nchroton.
[3] Also positrons were accelerated at the same energy. In general, we will refer to both just as electrons

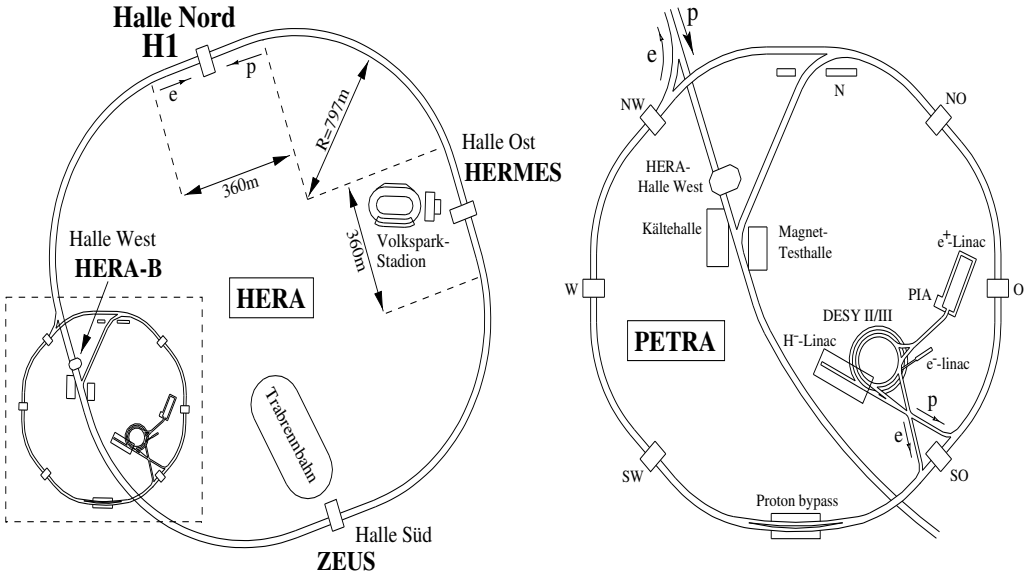


**Figure 3.1:** The HERA ring and the H1, ZEUS, HERMES and HERA-B experiments (left) and the DESY and PETRA facilities (right).

Electrons and protons were pre-accelerated at the DESY and PETRA rings and then injected into HERA with an energy of 12 GeV and 40 GeV, respectively. Superconducting magnets at 4.68 T and warm magnets at 0.165 T for protons and electrons, respectively, were used to bend the beams. Bunches from the two beams collided every 96 ns.

## 3.2 The H1 Detector

The H1 detector had three goals: a precise measurement of the scattered electron, hadron and jet detection and muon identification. For the electron identification, a spaghetti calorimeter (SpaCal) and liquid argon calorimeter (LAr) both with a fine granularity, were used. The hadronic part of the LAr provided also a good hadron and jet energy measurement. In addition, a tracking system in the inner part of the detector for track measurement and the return yoke for the magnetic field completed the H1 detector. The return yoke was instrumented and served as a muon detector as well as a device to measure hadronic energy leaking out the detector. In figure 3.2 an schematic view of the detector is shown.

In the detector design two other features were taken into account: the beam energy asymmetry and the high collision rates at HERA. The energy asymmetry, with a centre of mass boosted along the proton direction, required an asymmetric detector. Therefore, the detector in the proton direction (the forward direction) was more massive and segmented. The high collision rates and background level from non *ep* collisions had to be

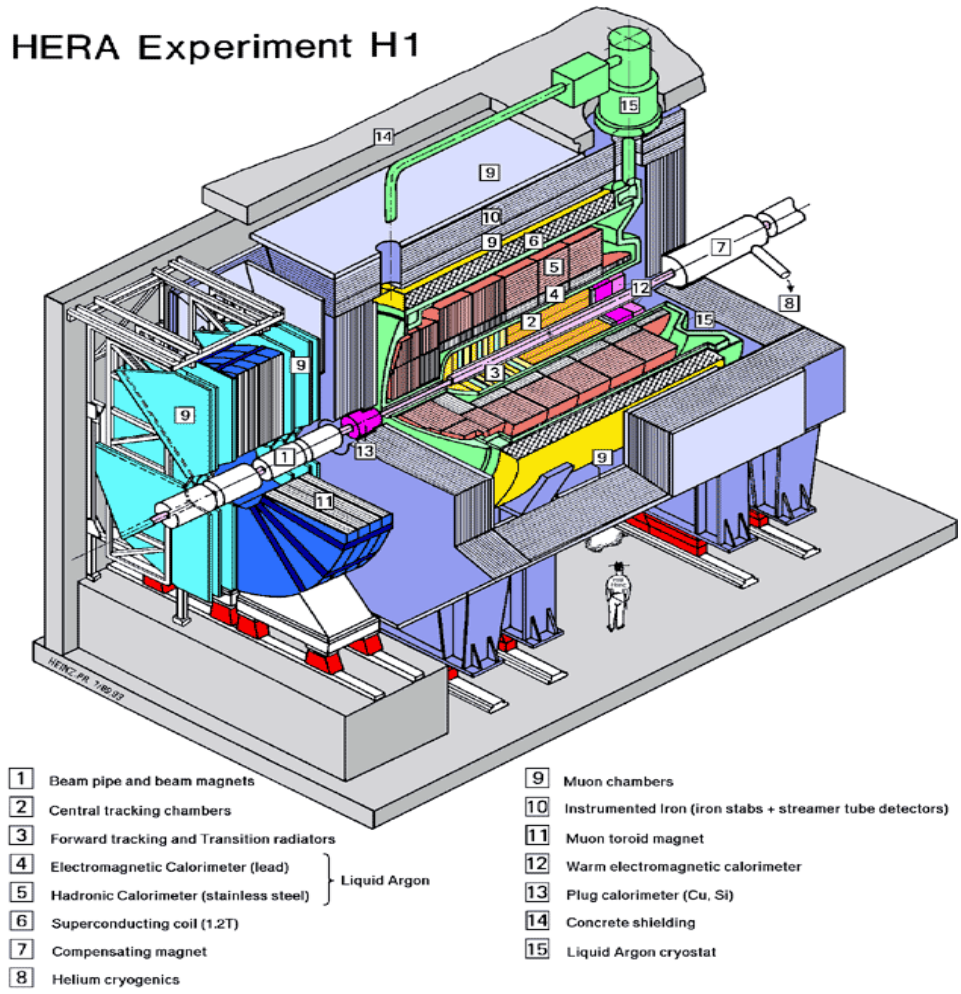


**Figure 3.2:** The H1 detector.

also considered in the trigger system, see section 3.2.4.

In this section a brief overview of the most relevant detector components for this analysis is given. For a more detailed description see [41] and [42].

### 3.2.1 Tracking System

The tracking system plays an important role since it delivers information for charged particle reconstruction and triggering. Due to the beam energy asymmetry, two tracking systems were build: the central and the forward tracking devices, see figure 3.5. In this analysis only the central tracking device (CTD) was used due to its better performance.

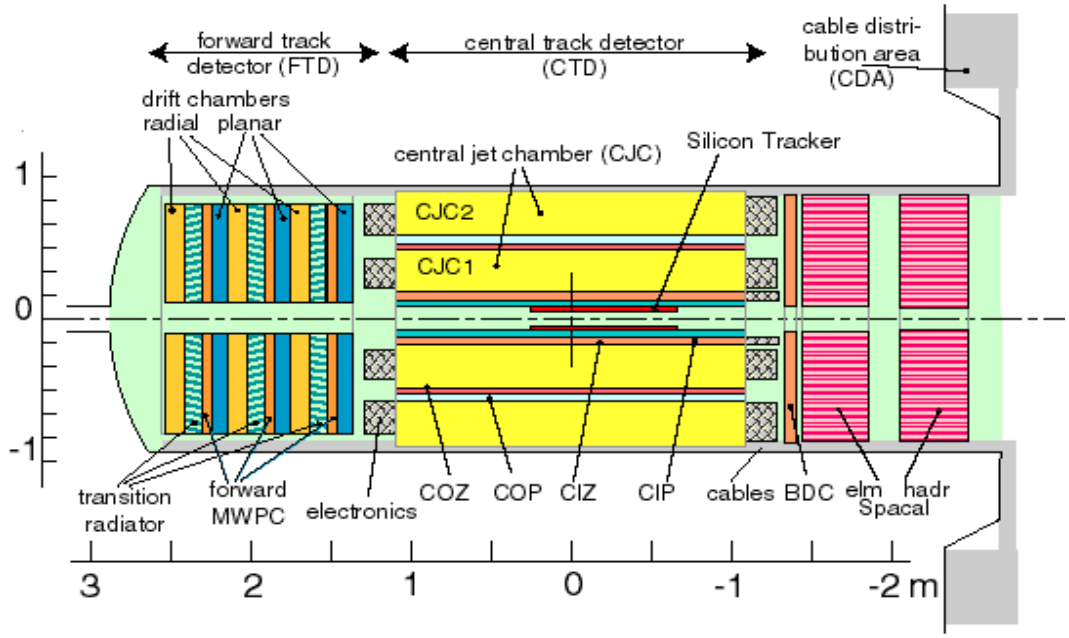


**Figure 3.3:** The tracking system at the H1 detector (HERA I)

From the beam pipe to the outer part of the detector it consisted of the Central Silicon Tracker (CST), the Central Inner Proportional Chamber (CIP), the Central Inner z-drift Chamber (CIZ), the Central Jet Chamber 1 (CJC1), the Central Outer z-drift Chamber (COZ), the Central Outer Proportional Chamber (COP) and the Central Jet Chamber 2 (CJC2).

The reconstruction in this region was based on the two concentric drift chambers CJC1 and CJC2 with their wires strung parallel to the beam. The CJC1, with an inner radius of 203 mm and outer radius of 451 mm, had 30 drift cells each with 24 sense wires and covered the polar angle angle $11° < \theta < 169°$. The CJC2, with an inner radius of 530 mm and outer radius of 844 mm, had 60 drift cells each with 32 sense wires and covered the polar angle range $26° < \theta < 154°$. The angular coverage of the CTD in this analysis is $25° < \theta < 160°$, $|\eta| < 1.5$.

The tracking system was inside a uniform magnetic field of 1.15 T parallel to the beams which is needed for momentum measurements. The drift angle of the charges liberated in the gas by an ionizing particle changes when applying a magnetic field. To compensate for this effect and therefore to obtain an optimal track resolution the drift cells were tilted 30° with respect to the radial direction. It forced ionization electrons to drift approximately perpendicular to high momentum tracks which led to a better spatial resolution. This had several additional advantages:

- Tracks crossed several cells. The wrong mirror track segments were then easily

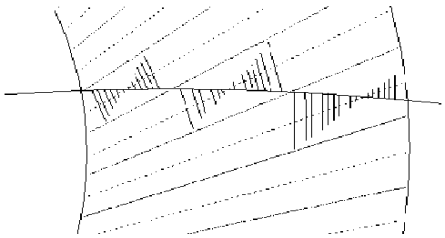

**Figure 3.4:** Track reconstruction in the CJC2. In the figure hits and mirror hits are shown. Mirror tracks are easy to identify since they do not continue into the neighbour cells.

determined since they did not continue in the neighbour cells and did not point to the event vertex, see figure 3.4.

- Tracks with $p_t > 400$ MeV crossed the sense wire plane in the CJC1 and CJC2 at least once. In these sense wires the drift time was shorter than 50ns which allowed to distinguish tracks from different bunch crossing.

The uniform magnetic field was produced by a superconducting solenoid which was outside the tracking and the calorimeter systems in order to avoid dead material in front of them. An instrumented iron yoke returned the magnetic flux and served as muon detector.

#### Central inner and outer z-drift chambers, CIZ and COZ

The jet chambers $r\phi$ resolution was $\sigma_{r\phi} = 170\mu$m and in $z$ $\sigma_z = 22$ mm. In order to improve the measurement in $z$ the CIZ and COZ chambers were used. Their wires were perpendicular to the beam axis. They achieved a resolution of $\sigma_z = 260\mu$m and $\sigma_z = 200\mu$m respectively.

#### Central inner and outer proportional chambers, CIP and COP

The CIP and COP proportional chambers were used only for a fast $z$ vertex position measurement. This was used in the level 1 trigger system (see section 3.2.4 and [43] [44]). The CIP and the CIZ chambers were replaced by the CIP 2000 chamber [45] during the detector upgrade.

#### Central silicon detector, CST

The tracking device closest to the beams was the silicon tracking detector (CST), [46] [47]. It was fully installed in the beginning of 1997. It was built to provide a precise vertex information by measuring charge particle tracks very close to the interaction point. It was

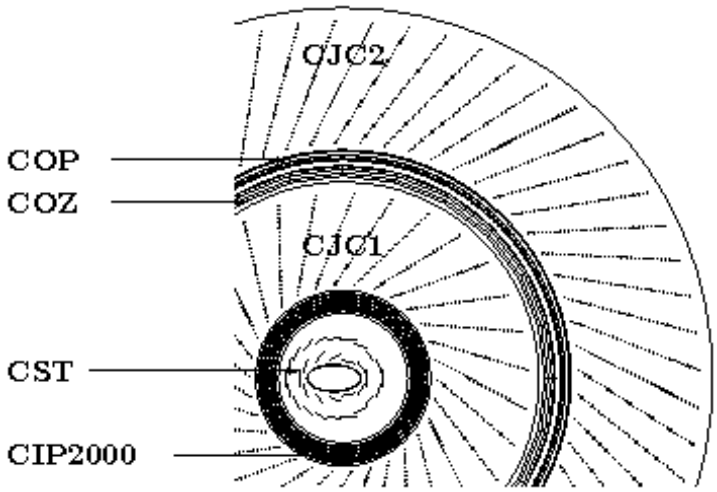


**Figure 3.5:** The tracking system at the H1 detector (HERA II)

composed of two concentric cylindrical layers of silicon sensors. These were 36cm long silicon strips, resulting in a polar angular coverage of $30° < \theta < 150°$. These two layers had 12 and 20 ladders at radii 5.75 cm and 9.75 cm respectively, see figure 3.6. After the detector upgrade in 2000-2001 the ladders were rearranged to accommodate a new elliptical beam pipe.

The resolution of the CST was $12\mu$m in the $r - \phi$ plane and $25\mu$m in $z$. The impact parameter resolution was $\sigma_{dca} = 37\mu$m for high momentum tracks and provided useful information in the determination of secondary vertices from B-mesons. The CST was also used to improve the track measurements from the CJC.

### 3.2.2 Calorimeters

The most important calorimeters for this analysis are presented in this section. The liquid argon calorimeter (LAr) was essential for the energy measurement of the scattered electron at high $Q^2 > 100$ GeV$^2$ values, the energy measurement of neutral and charged particles also at high particle densities, i.e. jets. Low energetic muons were measured in the LAr calorimeter but muons with an energy above 1.2 GeV only left a minimal ionizing signature and did not stop in the LAr. For this reason the muon system was built outside of the LAr, see figure 3.2. The LAr was situated inside the magnet while the muon system was situated outside. This choice minimized the dead material amount in front the LAr calorimeter. The magnet consisted of a superconducting coil and an iron yoke, the latter being instrumented with limited streamer tubes, LST, for the measurement of penetrating particles. Finally, the spaghetti calorimeter (SpaCal) covered the backward region (the

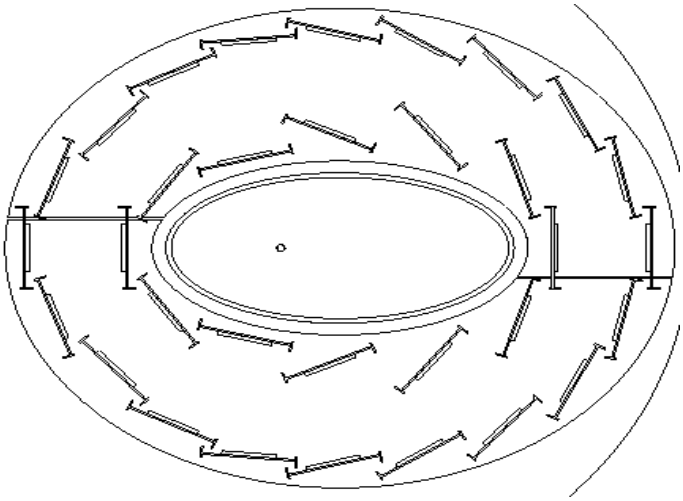

**Figure 3.6:** CST layers at radii 5.75 cm and 9.75 cm with 12 and 20 ladders respectively.

electron direction) and was specially important for electron measurement in DIS events.

## The Liquid Argon Calorimeter

The LAr calorimeter ( [48]) covered the polar angular range of $4° < \theta \lesssim 154°$, $-1.43 < \eta < 3.35$. As shown in figure 3.7 the LAr calorimeter was divided along the $z$ axis in eight self supporting "wheels" each one with a hadronic and an electromagnetic part, the latter in the inner side of the calorimeter.

It was a sampling calorimeter, i.e. it consisted on a passive or absorber material and a sampling or active material, the liquid argon. The absorber material were lead plate stacks in the electromagnetic part and stainless steel ones in the hadronic part. The orientation of the plate stacks was designed such that incident particles enter into the calorimeter as perpendicular to the plates as possible, typically angles larger than 45°, see figure 3.7. An incoming particle interacts with the absorber material producing secondary particles. If these have enough energy they start a cascade and thus produce showers of particles. Electrons and photons produce these showers via Bremsstrahlung and pair production and hadrons via elastic and inelastic scattering with nuclei. The number of particles in the shower is therefore proportional to the energy of the incoming particle. When these showers enter in the liquid argon, the charged particles ionize argon atoms and this signal is proportional to the energy of the incident particle.

The electromagnetic part consisted of 2.4 mm lead absorber stacks and 2.35 mm gaps between them of liquid argon. The hadronic one of 19 mm stainless steel and a double gap of 2.4 mm (4.8 mm of liquid argon). For the electromagnetic calorimeter there were 20 to 30 radiation lengths in the backward and forward regions, respectively, and for

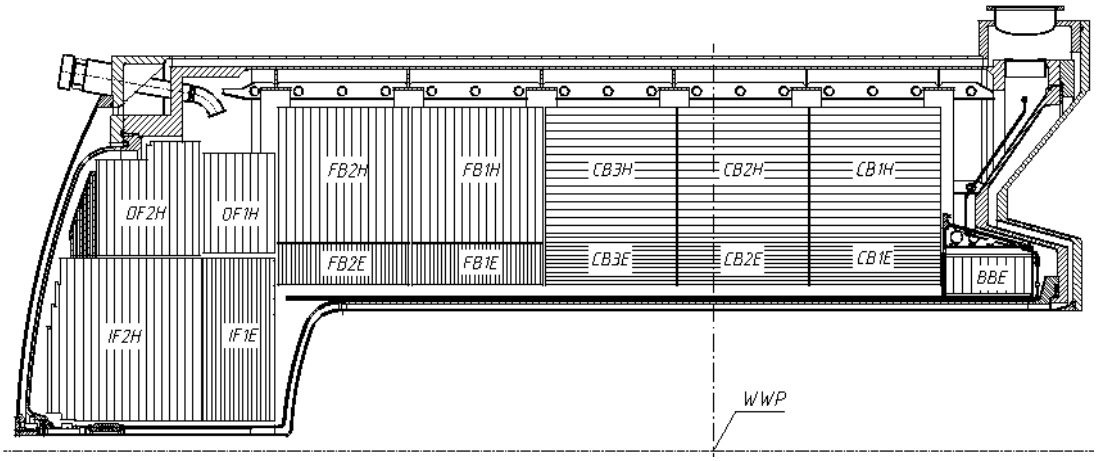


**Figure 3.7:** A longitudinal view of the liquid argon calorimeter, LAr.

the hadronic one about 5-8. The efficiency on collecting energy from electromagnetic particles in the LAr calorimeter was in average around 30% higher than for hadrons. This means, this energy imbalance had to be corrected during reconstruction, the LAr was a non-compensating calorimeter.

The LAr had more than 44000 readout channels, the calorimeter cells. The energy resolution for the electromagnetic and the hadronic calorimeters are respectively:

$$\frac{\sigma(E)}{E} = \frac{10\ldots13\%}{\sqrt{E/GeV}} + 1\% \tag{3.1}$$

$$\frac{\sigma(E)}{E} = \frac{50\%}{\sqrt{E/GeV}} + 2\% \tag{3.2}$$

## Spaghetti Calorimeter

The SpaCal [49] provided the essential electromagnetic as well as hadronic measurement in the polar angular range of $153° < \theta < 177.5°$. It was especially important for the electron identification at $1 < Q^2 < 100$ GeV$^2$ and therefore completing the $Q^2$ range of the LAr in the DIS regime. Outstanding for this analysis was its capacity for separating electrons and hadrons and therefore to identify fake electrons from photoproduction events, more abundant than those from DIS. At HERA most of the energy flow was in the proton direction. Therefore, the requirements for the hadronic energy measurement in the SpaCal were lower than those for the electromagnetic part.

The SpaCal had two independent sections: an inner one, the electromagnetic section, and an outer one, for hadronic energy measurement. The active material in the electromag-

netic part was 250 mm deep, which corresponds to 28 radiation lengths (the longitudinal leakage was negligible). For the hadronic section, the active material was also 250 mm deep, one radiation length. The resolution of electromagnetic and hadronic sections were $\sigma_{em}(E)/E \approx 7\%/\sqrt{E[GeV]}$ and $\sigma_{had}(E)/E \approx 30\%/\sqrt{E[GeV]}$, respectively.

### 3.2.3 Luminosity System

In order to calculate the luminosity, the rate of the Bethe-Heitler process ( [50]), $ep \to ep\gamma$, was used. This process is purely QED and does not depend on the inner structure of the proton. The most important background for this process is $eA \to eA\gamma$, where $A$ represents residual gas atoms in the beam and is expected to be 10% of the signal rate. This can be subtracted using the electron pilot bunches data.

To detect the scattered electrons and the outgoing photons the luminosity system used the electron tagger (ET), at $z = -33.4$m, and the photon detector (PD), at $z = -102.9$m with the requirement of coming from the same bunch crossing.

By using the ET and the PD quasi-real photoproduction events, $Q^2 < 0.01$ GeV$^2$, can be collected. This was done by applying independent thresholds on the electron, $0.2 < Ee'/E_e < 0.8$, and the photon energy, $0.004 < E_\gamma/E_e < 1.0$, and requiring the calibrated sum of both to be consistent, $E_\gamma/E_e = 1 - E_{e'}/E_e$.

### 3.2.4 Trigger System

At the HERA storage ring not only $ep$ collisions were observed. Background processes like proton-gas or proton-pipe scattering as well as synchrotron radiation from beam electrons were present. Beam halo and cosmic muons were a problem too. To disentangle these processes from actual $ep$ collisions and to store only interesting $ep$ collisions the trigger system was designed.

These background sources were not unique at HERA, they are common to other accelerators. Proton-gas collisions for instance are present even for a vacuum of about $10^{-9}$ hPa. The $off\ orbit$ protons hitting elements of the accelerator, the *beam wall background*, are present also in other accelerators like the TeVatron. However, the synchrotron radiation is important only when accelerating electrons and much less in hadron accelerators due to the mass of the particles. The time interval between two $ep$ bunch crossings was 96 ns, equivalent to 10.4 MHz, but actual $ep$ collisions toke place with a frequency of $\approx 100$ kHz. Thus non $ep$ collisions had a high frequency and a good efficiency rejecting background events was needed. Since it was not possible to have the detector readout in 96 ns because it would introduce large dead times other solutions had to be found.

Among the possible processes in $ep$ collisions the cross sections differences are large. While the photoproduction cross section is of the order of $\mu$b, the W production is of the order of pb. Since it was important to also collect those more rare events the trigger system

was also responsible to discriminate among them and avoid unnecessary dead time. This was done by down-scaling those events with higher cross sections, i.e. assigning weights.

In summary, the trigger system had to fulfil two targets. First, recognize *ep* collisions. Second, among those events recognize the process type and store them according its importance. The H1 trigger system had four levels. The time available to take a decision, the information and the precision from the detector components as well as the hardware used were different at each level.

## Level 1

Even very easy readout operations take longer than the 96ns between bunch crossings. Despite that level 1, L1, provided a trigger decision for each bunch crossing and therefore it was dead time free. To achieve this goal the system was pipelined with a decision delay of 2.5$\mu$s.

The smallest logical element were the trigger elements (TE). These TE were the output of nine subdetectors providing information for the decision taking. Examples of subdetectors are the vertex z-position system where the CIP and COP were used and the time of flight system which suppressed beam-wall and beam-gas events with the help of scintillators and the HERA clock. Some other subdetectors were based on information from the CJC, the LAr or the muon system.

The TE were organized in 128 subtriggers (ST) in total. In a ST the TE were combined using logical operators, i.e. AND ($\wedge$), OR ($\vee$) etc. The ST were designed to recognize not only *ep* events but also events with different cross sections, for instance photoproduction events, dijet events or DIS events. This was used to overcome dead time problems and to deliver an output rate of 1 kHz to the next trigger level since due to the large differences in cross sections for the observed processes prescales could be used for events with higher rates. This means that even if a ST decided to keep an event the trigger system kept the event according to the prescale. This is the difference between raw and actual trigger decisions.

An example of a subtrigger is the s83, used in this analysis (see 4.1). The subtrigger definition was designed to basically collect photoproduction events. Given the input rates no further restrictions were needed and the subtrigger level 2 and 3 were not needed.

## Level 2

If the event was kept by the L1 the pipelines were stopped and the readout was initiated. Thus the following bunch crossings were not considered and the dead time started. The L2 trigger level used 12 neural networks (L2NN) and topological triggers [51] (L2TT) to further select the events and reduce the input rates if needed. They provided 26 TE which validated the L1 decision. The time needed was typically 20$\mu$s and the input rate was reduced to 50 Hz. If the event was not taken the readout was stopped, the L1 pipelines

cleared and the dead time finished. In case the event was taken the readout continued and the event passed to the next level.

The L2 trigger level was not used in this analysis.

### Level 3

In the year 2006 the L3 trigger level was implemented using the Fast Track Trigger, (FTT) [52]. These tracks were used to cope with the higher background rates after the HERA luminosity upgrade and to identify more exclusive final states, specially in the heavy quark sector, for instance $D^*$ mesons in the *golden channel*, i.e. $D^* \rightarrow D^0\pi_s \rightarrow K\pi\pi_s$.

The L3 trigger level was not used in this analysis.

### Level 4

This level fully processed the incoming events, i.e. it made use of all the raw event information for the decision taking. Therefore L4 provided an event classification, a verification for the previous trigger levels as well as means to monitor them. This task was done by an asynchronous software and run by a computer farm.

At this level events were required to have either a hard scale, for instance high $Q^2$, $E_T^{miss}$, $E_T$, $p_T^{jet}$ or some track with $p_T > 1$ GeV, or to pass one of the specific final state finders, for instance $D^*$ or $J/\psi$ finders. In case they fullfiled at least one of these requirements the event is classified according to the L4 classes.

In order to check that these classification works not all events which did not pass any of those requirements were rejected. In order to ensure that no interesting physics is lost the non classified events were down-scaled, i.e. every $n$-th non classified event was recorded and a L4 weight was assigned to the event.

## 3.3 Offline reconstruction

At this level no event was rejected. Here a complete offline reconstruction with the H1REC package [53] and a calibration was done. The events used in this analysis were saved in a compressed format on the *data summary tape* (DST). The events which were assigned to a class were saved also on tape in the *physics output tape* (POT).

### 3.3.1 Track Reconstruction

The momentum of charged particles can be easily detected but the track reconstruction is done offline. Average number of charged particles is one of the observables used in

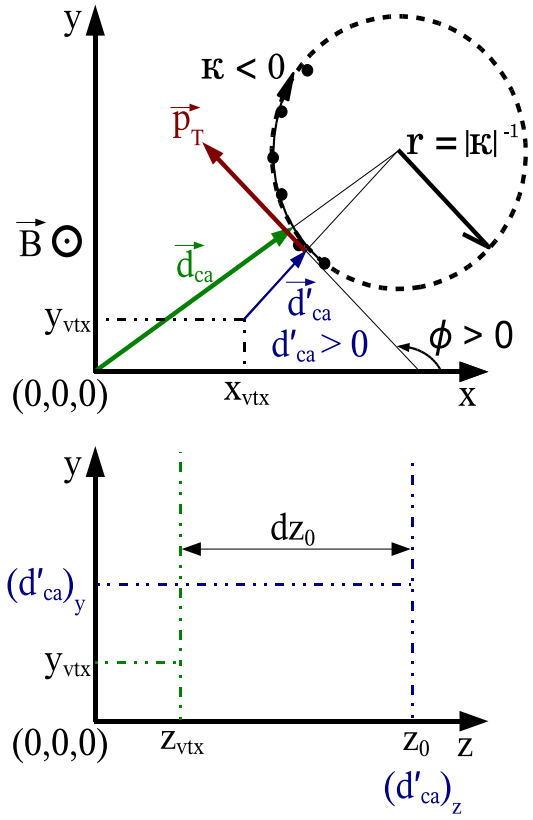


**Figure 3.8:** Helix parameters for a charged particle in a magnetic field perpendicular to the x-y plane and constant, $\vec{B} = B\hat{z}$. The upper figure corresponds to the detector transverse plane while the one below corresponds to the longitudinal section, x-y and y-z planes, respectively.

this analysis (section 2.4). The principle of track reconstruction is that under a uniform magnetic field parallel to the beams a charged particle move on a helical trajectory because of the Lorentz force $\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$, where in our case $\vec{E} = \vec{0}$, $\vec{B} = 1.15\hat{z}$ T (note this is in the proton direction) and $\vec{v}$ is the particle velocity. In order to parametrize the helical trajectory of a particle with momentum $\vec{p} = (\vec{p_t},\ p_z)$ we need five helix parameters, see figure 3.8:

- **Signed Curvature** $\kappa = \pm r^{-1}$ where $r$ is the radius of the trajectory in the x-y plane and it is positive if the particle movement is a counter-clockwise rotation as seen from the +z direction. The curvature is constant if the magnetic field $\vec{B}$ is perpendicular to the x-y plane and constant, $\vec{B} = B\hat{z}$, and the energy loss of the particle while moving through the medium can be neglected.

The curvature is related to the transverse momentum of the particle:

$$\kappa = \frac{zeB}{p_t} \rightarrow r = \frac{p_t}{zeB} \tag{3.3}$$

where $ze$ is the charge of the particle. For a particle with high transverse momentum the curvature is small and therefore the radius of the trajectory is large.

- **Signed Distance of Closest Approach** from the z-axis in the x-y plane, $d_{ca} = \pm|\vec{d_{ca}}|$. The sign is determined by $\frac{\vec{d_{ca}} \times \vec{p_t}}{|\vec{d_{ca}} \times \vec{p_t}|}$, i.e. $\vec{d_{ca}}$ and the trajectory direction form a right handed system.
- **Azimuth**, $\phi$, is the angle between the x-axis and $\vec{p_t}$ at the distance of closest approach.
- **Polar Angle**, $\theta$, is the angle between the momentum of the particle, $\vec{p}$, and the z-axis at the distance of closest approach.
- **z-position**, $z_0$, of the track at the point of closest approach

The first three parameters ($\kappa$, $d_{ca}$, $\phi$) can be determined by fitting a circle in the x-y plane.

Even if the track crossed the CJC1 and CJC2, where scattering between the charged particle and the material between the rings (2% $X_0$) was possible, it was usually enough to use a single parametrization set ($\kappa$, $d_{ca}$, $\phi$). In some cases two sets of parameters were needed. In this case the constraint that both join at the point between the two rings was applied.

The track parameters $d_{ca}$ and $z_0$ are calculated from the nominal interaction point. Another possibility is to calculate them from the actual interaction point. In this later case they are noted as $d'_{ca}$ and $dz_0$.

More details on track finding can be found in [42] and references therein.

After the track fit the momentum components of the charged particle were found with the equations 3.3 and:

$$\begin{aligned} p_x &= p \cdot sin\theta \cdot cos\phi \\ p_y &= p \cdot sin\theta \cdot sin\phi \\ p_z &= p \cdot cos\theta \end{aligned} \tag{3.4}$$

The relative momentum uncertainty $\sigma_p/p^2 < 0.01$ GeV$^{-1}$.

### 3.3.2 Analysis Framework

The present analysis was done within the Object Oriented framework [54] developed at H1 (H1OO) which is written in C++ and embedded in the ROOT framework [55]. ROOT provides generic tools for analysis and event display and a start up for more specific developments. The idea behind this project was to develop the specific tools needed for data analysis and store the relevant information from the event and particle reconstruction. These tools, classes, were developed by experts in a modular and extendable hierarchy and therefore easy to use for non-experts such that the analysis process becomes standardised. The classes were needed in order to cope with the event complexity at the H1 experiment due to high density of particles and the many and diverse topologies as well as the wide variety of analysis performed. In this way an overall high quality of data and analysis is ensured.

The data storage is organized in several layers: the Object Data Store (ODS) for the reconstruction output, the H1 Analysis Tag (HAT) for the basic event quantities and the micro Object Data Store (mODS) for particle information. Apart from these layers the user has the possibility to add her/his own layer with specific user information.

### 3.3.3 The Hadronic Final State

The hadronic final state (HFS) are all the particles except isolated electrons and muons. Within the H1OO framework a large amount of classes were developed to make the best use of the available information. Examples are particle finders like the electron, muon and the HFS or $D^*$ finders. In this section we discuss the HFS finder algorithm used in this analysis: Hadronic Reconstruction in OO (Hadroo2[4]) [56].

The proper reconstruction of the hadronic final state (HFS) plays an important role in this analysis. For this purpose information from the calorimeters and the tracking systems were combined in order to make the best profit of their information and improve the HFS measurement.

Briefly, the strategy used in Hadroo2 is to take the tracks and extrapolate them to the calorimeters. Tracks are then associated with some of the existing clusters in the calorimeters. The non paired clusters are supposed to come from neutral particles. From the paired tracks and clusters the uncertainty of both energy measurements are compared, i.e. $E_{track} \pm \sigma_{E_{track}}$ are compared to $E_{cluster} \pm \sigma_{E_{cluster}}$:

- If the track measurement is better: only the track measurement is used to define a charged particle. The energy amount corresponding to that of the track is removed from the cluster and if the remaining energy in the cluster is not compatible to zero that energy is associated to a neutral particle. In case it is compatible to zero the energy is considered as noise and removed.

[4] Hadroo2 is the successor of the previous Hadroo algorithm [57].

| Central | Combined |
|---|---|
| $P_T > 120$ MeV | $P_T > 120$ MeV |
| $20° < \theta < 160°$ | $0° < \theta < 40°$ |
| $\lvert d'_{ca} \rvert \leq 2$ cm | $\lvert d'_{ca} \rvert \leq 5$ cm |
| $R_{start} \leq 50$ cm | $R_{start} \leq 50$ cm |
| $R_{length} \geq 10$ cm if $\theta \leq 150°$ | $R_{length} \geq 0$ cm |
| $R_{length} \geq 5$ cm if $\theta > 150°$ | $\Delta p/p \leq 99999.9$ |
| $N_{CJChits} \geq 0$ | $N_{CJChits} \geq 0$ |
| | $\chi^2_{track-vertexfit} \leq 50$ |
| | $\chi^2_{cent.-fwd.tracker} \leq 50$ |

**Table 3.1:** "Lee West" track selection summary for central and combined tracks.

- If the cluster measurement is better: only the cluster measurement is used to define a charged particle.

In the following we discuss in more detail how this is done. First, we discuss about the input objects used by Hadroo2, tracks and clusters, and then how they are compared and manipulated in order to define combined objects, the HFS.

#### 3.3.3.1 Tracks

The tracks used by the Hadroo2 algorithm were measured in the central and in the forward tracking detectors. They were classified after the angular range of the tracks as central ($20° < \theta < 160°$), forward ($6° < \theta < 25°$) and combined ($0° < \theta < 40°$), see figure 3.9. Tracks which fulfil the selection in table 3.1 are called "Lee West" tracks[5] [58], where $R_{start}$ is the distance in the transverse plane from the interaction point to the nearest track hit (this condition ensures that the reconstructed track starts in the CJC1), $R_{length}$ is the track length, $N_{CJChits}$ the number of hits in the CJC and $\Delta p/p$ the relative track momentum uncertainty. In case a track fulfils more than one classification central tracks are preferred. In this analysis only the CTD was used and therefore no track classified as forward was selected. Only a small fraction of tracks were found to be classified as combined ($\ll 1\%$).

The tracks can be fitted to the primary vertex as well as to a secondary vertex[6]. The primary vertex classification is preferred.

[5] Named after Lee West who introduced this selection.

[6] For this purpose the silicon detector was used (section 3.2.1).

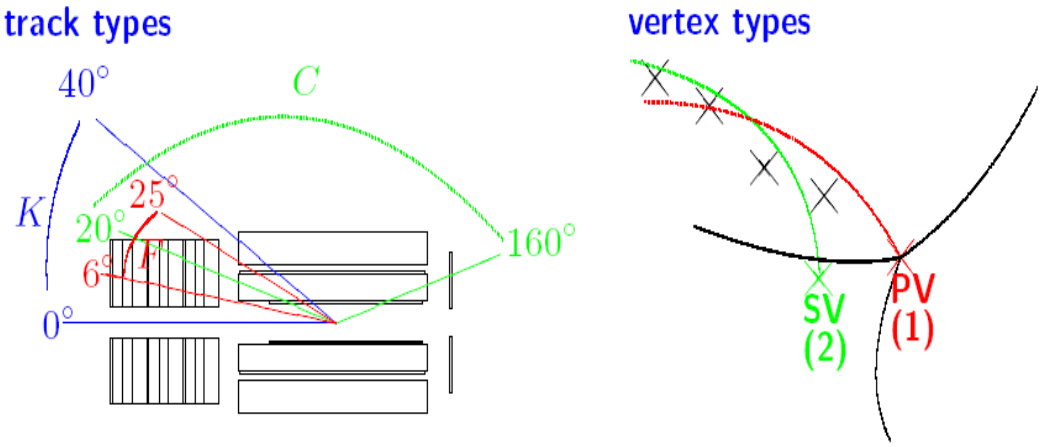


**Figure 3.9:** Different track types depending of their angular domain: central, forward and combined (left) and the two vertex hypothesis: primary vertex (PV) and secondary vertex (SV) (right).

#### 3.3.3.2 Clusters

The only clusters considered by Hadroo2 are the ones in LAr and SpaCal. The clusters four-momenta are reconstructed by adding massless cells and so they acquire a mass. The position of a cluster is determined by linear energy weighting of its cell positions.

As already mentioned in 3.2.2 the LAr was a non-compensating calorimeter. Thus, weighting algorithms were needed to take into account the lower response to hadrons compared to electromagnetic particles. This was done already before Hadroo2 by H1REC in the offline reconstruction. However, the definition of electromagnetic particles and hadrons is different here. In Hadroo2 a particle was considered to be an electromagnetic one if 95% of its energy was deposited in the electromagnetic part of the calorimeter and 50% of it in its first two layers. If this was not the case the particle was considered to be a hadron [59].

In Hadroo2 the noise in the calorimeters was considered. The origin of noise in the calorimeters were due to pile-up energy deposition from non-$ep$ collisions and noise in the electronics. Two types of algorithms were dedicated to it: topological background finders [41], dedicated to the search of non-$ep$ collision and therefore for event rejection, and noise in the $ep$ events cell finders, where only noise cells were removed while the event was kept [56]. An example of the latter are clusters where $E_{clu} < 0.2$ GeV in the LAr or $E_{clu} < 0.1$ GeV in the SpaCal, which were removed from the event.

#### 3.3.3.3 Tracker and Calorimeter Measurement Comparison

The energy uncertainty of a track can be obtained by error propagation:

$$E^2_{track} = \vec{P}^2_{track} + m^2_\pi = \frac{\vec{P}^2_{T,\ track}}{sin^2\theta} + m^2_\pi \rightarrow \frac{\sigma_{E_{track}}}{E_{track}} = \frac{1}{E_{track}}\sqrt{\frac{\vec{P}^2_{T,\ track}}{sin^4\theta} cos^2\theta\ \sigma^2_\theta + \frac{\sigma^2_{P_T}}{sin^2\theta}}, \tag{3.5}$$

where every track is supposed to be originated from a pion, $\sigma_{P_T}$ and $\sigma_\theta$ are the uncertainties on $P_{T,\ track}$ and the polar angle of the track, $\theta$, respectively. Correlations between these two variables are neglected. However, taking the corrections into consideration by using the full covariance matrix gives same the results within 2%.

The paired cluster energy to a track can receive contributions from neutral particles. This contribution is not known and therefore the relative error of the calorimetric measurement cannot be done alone with the information from the calorimeter. This means that the relative error of the cluster energy has to be track based. In the LAr the following assumption [41] is made:

$$\frac{\sigma_{E_{LAr}}}{E_{LAr}} \equiv \frac{\sigma_{E_{LAr}}}{E_{track}} = \frac{0.5}{\sqrt{E_{track}}}. \tag{3.6}$$

Following 3.5 and 3.6 a track is considered as good if:

$$\frac{\sigma_{E_{track}}}{E_{track}} < \frac{\sigma_{E_{LAr}}}{E_{track}} \tag{3.7}$$

From equations 3.5 and 3.6 we see that tracks are better measured at low $P$ (most of tracks with $P \lesssim 20$ GeV are considered better than the calorimetric measurement). For this reason, low $P_T$ tracks are treated first where the track measurement is better and thus the calorimetric cluster information is not used.

If the energy of the clusters

$$E_{clusters} < E_{track} \times \left[1 + 1.96\sqrt{\left(\frac{\sigma_{E_{track}}}{E_{track}}\right)^2 + \left(\frac{\sigma_{E_{LAr}}}{E_{LAr}}\right)^2}\right] \tag{3.8}$$

and condition 3.7 fulfil the clusters are removed from the list. Otherwise, only the amount of energy $E_{track}$ is removed. This is done cluster after cluster by increasing distance from the track extrapolation to the calorimeter.

If the calorimetric measurement is better, i.e condition 3.7 and

$$E_{track} \in [E_{clusters} - 1.96\sigma_{E_{clusters}}, + \infty] \tag{3.9}$$

only the calorimetric measurement is used to define a particle candidate. If the condition 3.9 is not fulfilled the track measurement is used and its energy subtracted from the calorimetric energy cluster after cluster by increasing distance from the track extrapolation to the calorimeter.

After all tracks have been considered, particle candidates corresponding to neutral hadrons are made out of the remaining clusters.

## 3.4 Calibration of Jet Energy Measurement

In hadronic showers a non negligible amount of energy cannot be measured: the recoil of the atoms in the absorber plates, muons from meson decays and neutrinos are not measured. As a consequence, the measured energy is smaller than the energy carried by the hadron. The HFS definition depends on the uncertainty of the energy measurement and a calibration of the calorimeters is performed in order to improve it.

Using the high granularity of the LAr a software is used to weight the clusters response and compensate for losses.

The method used by this analysis exploits the fact that since protons and electrons collide head-on the transverse momentum of the final state has to be zero, i.e. the transverse momentum of the scattered electron balances the transverse momentum of the HFS and isolated electrons and muons. The method uses DIS data where the electron is detected in the SpaCal and the electron is measured with a high precision.

The electron measurement in the SpaCal is used to calculate its transverse momentum $P_T^e$ while for the hadronic system $P_T^h = \sqrt{(\sum P_x^i)^2 + (\sum P_y^i)^2}$ is used. To study how the transverse momentum is balanced, the quantity $P_T^{bal} = \frac{P_T^h}{P_T^e}$ is used. In order to study how well the detector simulation and reconstruction is performed the double ratio $DR(P_T^{bal}) = \langle \frac{P_T^h}{P_T^e} \rangle_{data} / \langle \frac{P_T^h}{P_T^e} \rangle_{MC}$ is defined. Using these quantities, it can be seen that the energy can be measured with an uncertainty of less than 2%. Details on the calibration can be found in [60].

# Chapter 4

# Event and Signal Selection

In this chapter the event selection for this analysis is described. The data were recorded in the years 1999 and 2000 by the H1 detector. In that period the centre-of-mass energy was $\sqrt{s} \simeq 318$ GeV. The corresponding integrated luminosity is 48 $pb^{-1}$. The selection consists of tagged photoproduction events with dijets. With these events two analyses are performed: a charged particle and a low $P_T$ jet, the so called minijet, multiplicity analysis. The same subtriggers were used for the two analyses since both require photoproduction events.

## 4.1 Online Event Selection

The online selection is the event selection done by the subtriggers (ST) in the years 1999 and 2000. Only those events where the important detector components for the analysis were fully operational are included. Those are the luminosity system for tagged photoproduction detection, the jet chambers (CJC1 and CJC2), the liquid argon calorimeter (LAr), the central silicon tracking detector, the inner and outer proportional and z-drift chambers (CST, CIP, COP, CIZ and COZ) for track and dijet measurements, the spaghetti calorimeter (SpaCal) to exclude DIS events, and the time of flight (TOF) for background rejection. In addition, only runs classified as good or medium[1] and with a luminosity larger than 0.1 $\text{nb}^{-1}$ are considered. Minimum bias runs are not included.

At H1 not only proton-electron collisions were observed and among $ep$ collision cross sections the differences are of several orders of magnitude. For this reason, a subtrigger strategy had to be designed in order to select interesting events. In our case these were photoproduction events and only the level 1 (L1) and level 4 (L4) were used.

[1] Runs were classified as good, medium or poor depending on the run conditions, like background and whether important subsystems were working (tracking system, LAr, etc).

| Subtrigger | Definition |
|---|---|
| s83 | (LU_ET) $\wedge$ $\neg$(LU_PD_low) $\wedge$ (DCRPh_Tc) $\wedge$ (zVtx_sig $>$ 1) |

| Trigger Element (TE) | Short description |
|---|---|
| LU_ET | energy deposition in the ET33 |
| LU_PD_low | energy deposition in the photon detector |
| DCRPh_Tc | at least three fired track masks with $P_T \geq 450$ MeV |
| zVtx_sig | significant maximum in the z-vertex histogram (figure 4.1) |

**Table 4.1:** Subtrigger s83 definition and trigger elements description.

### 4.1.1 L1 Trigger Selection

In this section the ST 83 (s83) and its trigger elements (TE) are described. First, a short description of the TE is given in order to better understand its function. Then, a more detailed definition is given.

For the tagged photoproduction selection the ST 83 (s83) was used. This ST consists of basically four trigger elements (TE), see table 4.1. The LU_ET TE ensures that the scattered electron was found in the electron tagger, ET33. The $\neg$(LU_PD_low) TE[2] selects events in which the energy deposition in the photon detector (PD) is smaller than the deposition in the ET33 and therefore ensures photoproduction events. The DCRPh_Tc TE requires at least three track masks from the interaction point to fire. This rejects beam induced background, cosmic muons and synchrotron radiation. Finally, the zVtx_sig $>$ 1 TE ensures that an event vertex was reconstructed close to the nominal interaction point and further reduces background.

#### Luminosity system: LU_ET and LU_PD_low

From the luminosity system (section 3.2.3) two TE are used: LU_ET and LU_PD_low. With the LU_ET condition an energy deposition in the ET33 above 4 GeV is required and as a result events with $Q^2 < 0.01$ GeV$^2$ are automatically selected. Besides it, the TE $\neg$(LU_PD_low) condition requires that the energy deposition in the PD is below a given threshold. All together we have a veto on Bethe-Heitler processes, $ep \rightarrow ep\gamma$, and trigger on events with a scattered electron in the ET33 for photoproduction event selection.

#### Central jet chamber trigger: DCRPh_Tc

The DCRPh trigger [61,62] used information from the central jet chamber trigger to find

[2]The symbol $\wedge$ represents a logical AND and $\neg$ a logical NOT.

tracks. For this purpose 10 selected radial wire layers out of 56 were used. From the information delivered by these layers, hit masks were defined in drift space and curvature. 10000 different masks were then used to recognize the tracks. With this method transverse momentum track recognition as well as the charge for low momentum tracks ($< 1$ GeV) and three different thresholds on the total number of tracks could be achieved. These three thresholds are the TE: DCRPh_Ta, DCRPh_Tb and DCRPh_Tc for one, two and three tracks, respectively.

With the DCRPh_Tc TE condition we ensure that at least three track masks with $p_T \geq 450$ MeV fired with a distance of closest approach to the nominal interaction point of less than 2 cm and therefore we suppress beam-wall events as well as synchrotron radiation background and cosmic muons.

#### z-vertex trigger: zVtx_sig

The purpose of the z-vertex trigger is to estimate online the event vertex position along the beam axis [63, 44]. Only event vertices inside the nominal *ep* interaction region are accepted since events outside this region are interpreted as background processes.

The signals from the central proportional chambers CIZ and COZ (section 3.2.1) and the first forward proportional chambers (FPC) were used. The vertex estimator combines the four cathode pad signals to build rays as shown in figure 4.1 (straight lines along these signals pointing the z-axis). These rays are then filled in a histogram with a 16 fold segmentation in $\phi$ sectors and 16 bins along the z-coordinate with a bin width of 5.4 cm. Since the combinatorial (wrong) entries of rays are supposed to be randomly distributed the bin with the larger number of entries is expected to be filled with rays with correct combinations of pad signals. If the event vertex is outside the nominal interaction region the histogram contains therefore background rays and does not develop significant peaks.

The z-vertex trigger has different significant thresholds for the peak. The condition zVtx_sig $> 1$ ensures a significant peak close to the nominal interaction region.

### 4.1.2 L4 Trigger Selection

Most of the events were accepted at this level by different event classes, for instance as jet or multi-particle classes. If no hard scale was present in an event it was down-scaled (section 3.2.4). After the offline event selection (section 4.2) 1.9% of the events were down-scaled with a weight of 10. All these events were taken with the corresponding L4 weights. About 65% of the weighted events were found to be dijet events with the leading jet transverse momentum in the range $5 < P_T^{Jet} < 7$ GeV.

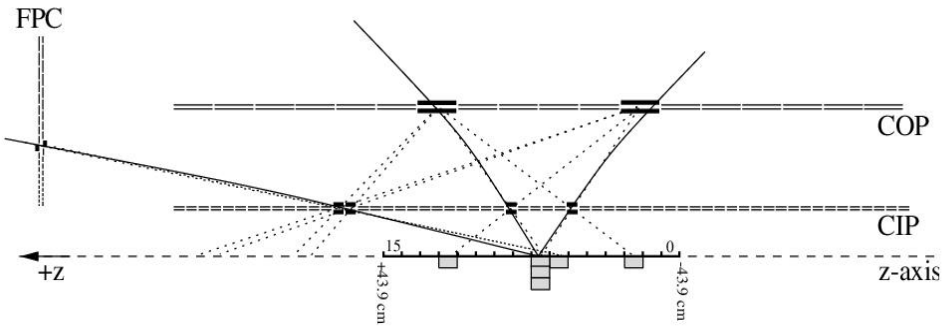


**Figure 4.1:** The vertex estimator combines four cathode pad signals to build rays. The combinatorial wrong entries (dotted lines) of rays are supposed to be randomly distributed. The correct ones (full lines) point to the event vertex.

## 4.2 Offline Event Selection

In the online event selection strategies to detect and store specific processes were used. In this analysis we are interested in photoproduction events and therefore the s83 was used (section 4.1.1). Because of the detector dead time, fast decisions had to be made while keeping a high acceptance and a high background rejection efficiency. Nevertheless, not all events selected by the s83 are of interest in this analysis since for example the s83 does not require any jet. In the offline selection the selection criteria is tightened. The advantage here is that all information from the detector can be used without time constrains.

### 4.2.1 Kinematic Selection

There are different methods to reconstruct the event kinematics, i.e. the photon virtuality, $Q^2$, and the inelasticity, $y$. Since the electron measurement in the ET33 is always required it can be used to determine $Q^2$ and $y$. With the electron energy before and after scattering, $E_e = 27.6$ GeV and $E_{e'}$, respectively, and the electron scattering angle, $\theta_{e'}$ they are calculated as:

$$Q^2 = 4E_eE_{e'}\ cos^2\frac{\theta_{e'}}{2} \tag{4.1}$$

$$y = 1 - \frac{E_{e'}}{E_e}\ sin^2\frac{\theta_{e'}}{2} \approx 1 - \frac{E_{e'}}{E_e} \tag{4.2}$$

Due to the low angle of the ET33a tagged electron in the ET33 corresponds to quasi-real photoproduction event selection, $Q^2 < 0.01$ GeV$^2$.

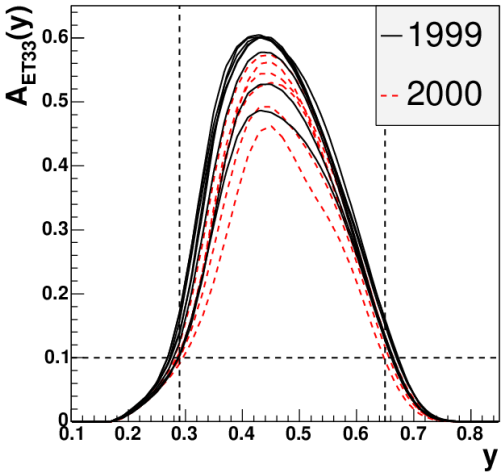


**Figure 4.2:** Electron tagger acceptance for different years and their run periods as a function of the inelasticity, y. In the range $0.3 < y < 0.65$ an acceptance above 10% is achieved.

| Variable | Selected range | Description |
|---|---|---|
| $y$ | $0.3 \ldots 0.65$ | High ET33 acceptance |
| $\lvert x_{ET33} \rvert$ | $< 6.5$ cm | Deposition well inside the ET33 ($Q^2 < 0.01$ GeV$^2$) |

**Table 4.2:** Summary of the kinematic selection.

The ET33 acceptance is not constant over the inelasticity, $y$, and is highly dependent on the exact beam parameters as shown in figure 4.2 [64]. A reasonable acceptance, above 10%, is reached in the inelasticity range $0.3 < y < 0.65$. With this requirement, the average ET33 acceptance amounts 40%. The limited acceptance is included when correcting for trigger and detector effects.

A good electron energy measurement requires the electron energy to be fully deposited in the ET33. This is achieved if the deposition of the energy in the ET33 is well inside the ET33 by requiring that the x-coordinate of the deposition fulfils $|x_{ET33}| < 6.5$ cm [64].

Table 4.2 summarizes the ET33 and kinematic selection.

### 4.2.2 Jet Selection

In chapter 1 we saw that by requiring jets at large $P_T$ a hard scale needed for the application of perturbative QCD is provided. In 1.7 the inclusive $k_t$ *algorithm* was presented. This jet finder is used in this analysis in the $p_t$ *weighted recombination scheme*. In 3.3.3

| Charged particle multiplicity | Mini-jet multiplicity |
|---|---|
| $P_T^{Jets} > 5$ GeV<br>$-1.5 < \eta^{Jets} < 1.5$ | $P_T^{Jets} > 5$ GeV<br>$-1.5 < \eta^{Jets} < 2.79$ |

**Table 4.3:** Summary of the event jet selection for the charged particle and the minijet multiplicity analysis.

it was shown how the track and cluster information is merged to define the hadronic final state, HFS. The HFS is the input for the $k_t$ *algorithm*. In this analysis, events with two jets were selected by requiring[3] $P_T^{Jets} > 5$ GeV.

For the **charged particle multiplicity studies** the pseudorapidity region $-1.5 < \eta^{Jets} < 1.5$ is chosen. This corresponds to the acceptance of the central tracking device (CTD) where the charged particles are measured (section 4.2.4). The reason for this choice is to have the jets and the charged particles measurement in the same $\eta$ range. The leading jet, the jet with highest $P_T$, defines the toward region (see 2.4).The toward region is defined to contain the charged particles from the leading jet and therefore it must be in the $\eta$ range used for the track measurement.

For the **minijet multiplicity studies** the pseudorapidity region $-1.5 < \eta^{Jets} < 2.79$ is chosen. This ensures that the jets are well inside the liquid argon acceptance.

Table 4.3 summarizes the jet selection.

### 4.2.3 Trigger Efficiency (s83)

The efficiency of a subtrigger is the probability that a signal event is detected by the subtrigger: $\epsilon_{ST} = N_{ST}/N_{total}$, where $N_{ST}$ is the total number of events detected by the ST and $N_{total}$ is the total number of signal events in which we are interested. In this analysis the ST is s83, $\epsilon_{ST} = N_{s83}/N_{total}$.

Since it is not possible to know directly the number of non detected events the efficiency has to be estimated by using monitor subtriggers (MST). The idea is to select a set of events where all cuts were applied but using an orthogonal ST with low thresholds, i.e. the MST and therefore such that events were randomly collected with respect to s83. From these events and the ones selected by s83, the s83 efficiency can determined.

[3] Although the $P_T^{Jets}$ region of study in this analysis is $P_T^{Jets} > 5$ GeV we select events with dijets with $P_T^{Jets} > 3.5$ GeV. Those events where $P_T^{Jets} > 3.5$ GeV and at least one jet is in the range $3.5 < P_T < 5$ GeV were also used in the unfolding procedure (section 5.3).

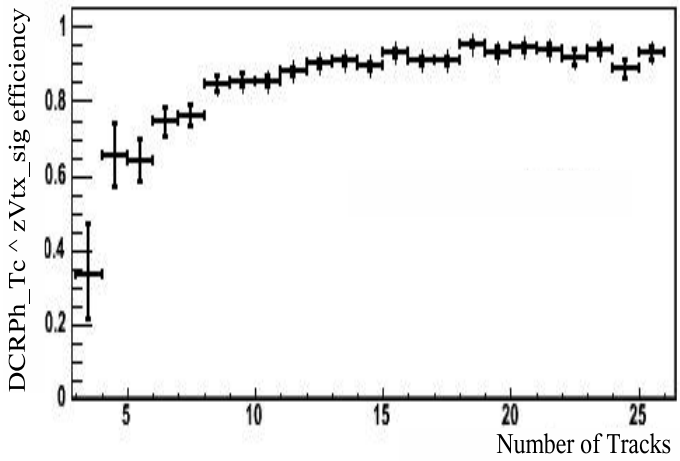


**Figure 4.3:** DCRPh_Tc and zVtx_sig TE efficiency as a function of the number of tracks. For track multiplicities larger than 10 the efficiency is large and constant while it drops at lower track multiplicity values.

Thus, $N_{total} = N_{MST}$ and $N_{s83} = N_{MST \wedge s83}$. The efficiency is calculated as:

$$\epsilon_{ST} = \frac{N_{MST \wedge ST}}{N_{MST}}. \tag{4.3}$$

In the s83 definition we have two TE for tagged photoproduction event selection: LU_ET and $\neg$(LU_PD_low). In section 4.2.1 the acceptance for these TE was shown and it is taken into account when correcting for trigger and detector effects.

For the TE DCRPh_Tc and zVtx_sig the monitor triggers s0 and s3 were used as MST. These MST were used to select deep inelastic events (DIS) and were based in SpaCal conditions which were not present in the s83 definition. Therefore, they are considered to be independent. For efficiency studies dijet events (table 4.3) in DIS, i.e. triggered by s0 and s3 were considered.

Since a high number of tracks is expected to influence the L1 vertex estimator, zVtx_sig, it was important to study the efficiency as a function the number of tracks. This is shown in figure 4.3 (only statistical uncertainties are shown). We see that for large track multiplicities ($\gtrsim 10$) the efficiency is close to one while at low track multiplicities the efficiency drops.

In order to see whether this could introduce a bias for the analysis, the relative appearance of events with low track multiplicity was studied. This is shown in figure 4.4. The mean value of the distribution is around 20 (maximum at 16) and the relative appearance of events with a track multiplicity lower than 10 is low. Therefore, the effect of the rather low efficiency at low track multiplicity is very small.

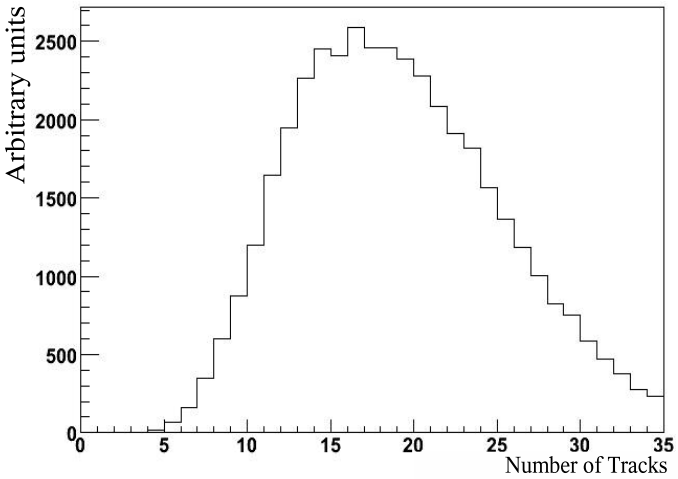


**Figure 4.4:** Event distribution as a function of the number of tracks. The mean number of tracks per event is around 20. Events with less than 10 tracks have a low efficiency but have a low relative appearance.

This was explicitly checked in the worse case scenario for low track multiplicity events by considering the efficiency for track multiplicities of 3, 4 and 5 the values from the lower error bands when correcting data from efficiencies[4]. The effect can be seen in figure 4.5. Here, the average value of charged particles in the transverse regions is compared when no track multiplicity efficiency is considered (solid line) and when the efficiency is considered (taking the lower bounds given by the uncertainties in the first three bins). Although the effect is small the track efficiency is taken into consideration for further calculations (without modification).

### 4.2.4 Track Selection

The tracks considered in this analysis are tracks coming from the event vertex. Tracks from secondary vertices are not considered since the origin of those particles is related to other processes like nuclear interactions between particles from the *ep* collision and detector components. Therefore, charged particles from weak decays are also excluded.

The TE DCRPh_Tc at the level 1 in the trigger system requires three track masks with $P_T > 450$ MeV (see 4.1.1). However, after the offline reconstruction where all information from the detector was used a better track reconstruction was achieved. For consistency, the events were required to have at least three primary vertex fitted tracks with $P_T > 450$ MeV both in data and the Monte Carlo detector simulations.

In the case of "Lee West" tracks, section 3.3.3.1, a selection on the tracks was applied to

[4]The reason for it are the quite large uncertainties in these bins.

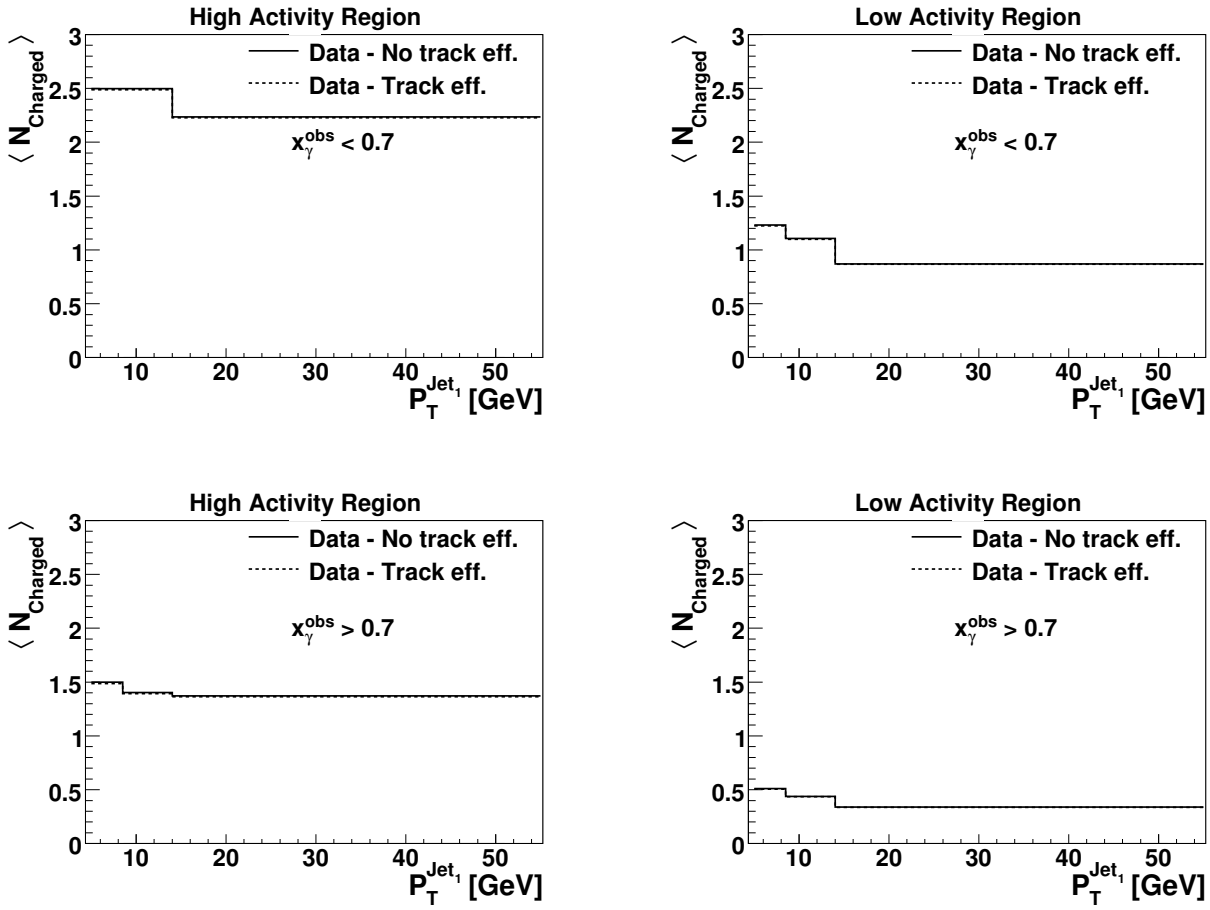


**Figure 4.5:** The average value of charged particles in the transverse regions is compared when no track multiplicity efficiency is considered (solid line) and when the efficiency is considered and taking the lower bounds given by the uncertainties in the first three bins (dashed line).

ensure good track quality. Following similar criteria, in this analysis a harder selection over the "Lee West" tracks was applied. The track selection summary is shown in table 4.4.

| Primary vertex fitted tracks (at least 3 with $P_T \geq 450$ MeV) | Short description |
|---|---|
| $P_T > 150$ MeV | Minimum $P_T$ track[5] |
| $-1.5 < \eta < 1.5$ | Pseudorapidity range covered by the CTD |
| $\lvert d'_{ca}\ sin\theta \rvert \leq 1$ cm | Distance of closest approach from the actual event vertex |
| $R_{start} \leq 30$ cm | Track start position in the x-y plane[6] |
| $R_{length} \geq 10$ cm | Minimum track length |

**Table 4.4:** Analysis track selection summary.

## 4.3 Event and Control Plots for Charged Particle Analysis

Control distributions for the event selection are shown in figures 4.6 and 4.7. The distributions are normalized to the area. The data are compared to the PYTHIA Monte Carlo where detector effects were simulated and reconstructed. Only statistical uncertainties in the data are shown (they are negligible for the Monte Carlo). The energy fraction of the photon entering the hard scattering $x_\gamma^{obs}$ (left) and the photon proton centre-of-mass energy $W_{\gamma p}$ (right) distributions are shown in figure 4.6.

The variable $x_\gamma^{obs}$ is used to select events with resolved or point-like photons (section 2.4), where events with $x_\gamma^{obs} < 0.7$ are considered to be resolved and $x_\gamma^{obs} > 0.7$ to be direct. The discrepancies seen in the $x_\gamma^{obs}$ distribution are not important since detector effects are treated explicitly by the unfolding method (chapter 5).

In figure 4.7 several distributions for the leading jets are shown: the transverse momentum distribution of the leading jet $P_T^{Jet}$ (also in logarithmic scale), the polar angle distribution of the leading jet $\theta^{Jet}$, the azimuthal angle distance of the second leading jet with respect of the leading jet $\Delta\phi(Jet_1 - Jet_2)$ and the $\Delta R(Jet_1 - Jet_2)$[7].

The discrepancies observed in the transverse momentum of the leading jet, $P_T^{Jet}$, are also explicitly treated by the unfolding method. The $\Delta\phi(Jet_1 - Jet_2)$ discrepancies were studied when unfolding data. They were found not to modify the unfolded distributions.

[5] A minimum track $P_T$ and length is needed in order to measure the helix parameters properly, see 3.3.1.

[6] With this selection it is ensured that the track starts in the CJC1, i.e. to recognize neutral particles undergoing nuclear interactions while crossing the dead material between the CJC1 and CJC2.

[7] $\Delta R(Jet_1 - Jet_2)$ is defined as $\sqrt{(\phi^{Jet_1} - \phi^{Jet_2})^2 + (\eta^{Jet_1} - \eta^{Jet_2})^2}$.

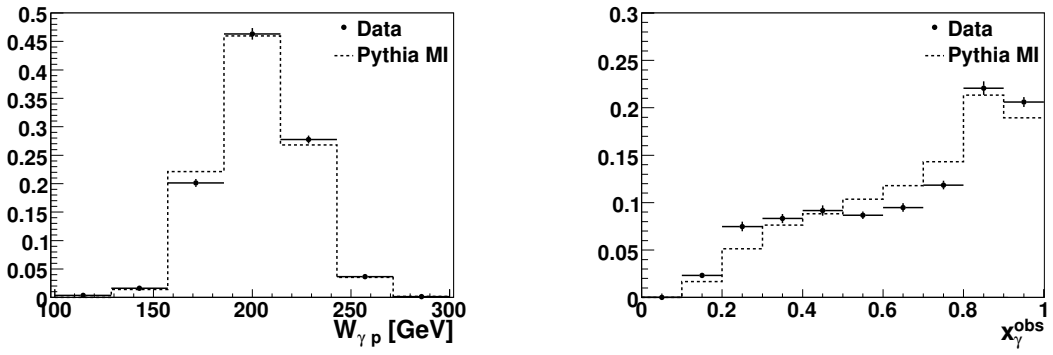


**Figure 4.6:** Energy fraction of the photon entering the hard scattering $x_\gamma^{obs}$ (right) and the photon proton centre-of-mass energy $W_{\gamma p}$ (left) control distributions after event selection for the charged particle studies.

In figures 4.8 and 4.9 control plots are shown for track selection and detector track simulation and reconstruction, respectively. The distributions are normalized to the area and only statistical uncertainties for the data are shown (for the Monte Carlo they are negligible).

In figure 4.8 the transverse momentum $P_t^{track}$ (also in logarithmic scale), pseudorapidity $\eta^{tracks}$, the azimuthal angle $\phi^{track}$ and the $\Delta R(Jet_1 - tracks)$ distributions for the charged particles are presented. The discrepancies in the $\eta^{tracks}$ distribution at $\eta^{tracks} \simeq 1$ were studied using different detector simulations, i.e. different dead material descriptions, and looking at different $\phi_{tracks}$ and $P_t^{track}$ ranges. No reason to doubt about the detector simulation and reconstruction was found and therefore neither about the unfolding.

The track length and the logarithm of the actual distance of closest approach distributions are shown in figure 4.9. Most of the tracks have a length much longer than 10 cm and therefore the discrepancies around 20 and 60 cm, where the Monte Carlo description overestimates and underestimates the date, respectively, are not important, i.e the track selection is not affected by the $R_{length} > 10$ cm condition. Only the tracks with $\log dca' > 0$ are affected by the $|dca' sin\theta| < 1$ cm condition (depending on $\sin\theta$ these tracks are accepted or rejected). Tracks where $\log dca' < 0$ are always accepted and therefore the discrepancies around -4 and -1, where the Monte Carlo description underestimates and overestimates the data, respectively, are not important.

## 4.4 Minijet Control Plots

Control distributions for the event selection are shown in figures 4.10 and 4.11, where all distributions are normalized to the area and only statistical uncertainties shown. The data are compared to PYTHIA where detector effects are simulated and reconstructed.

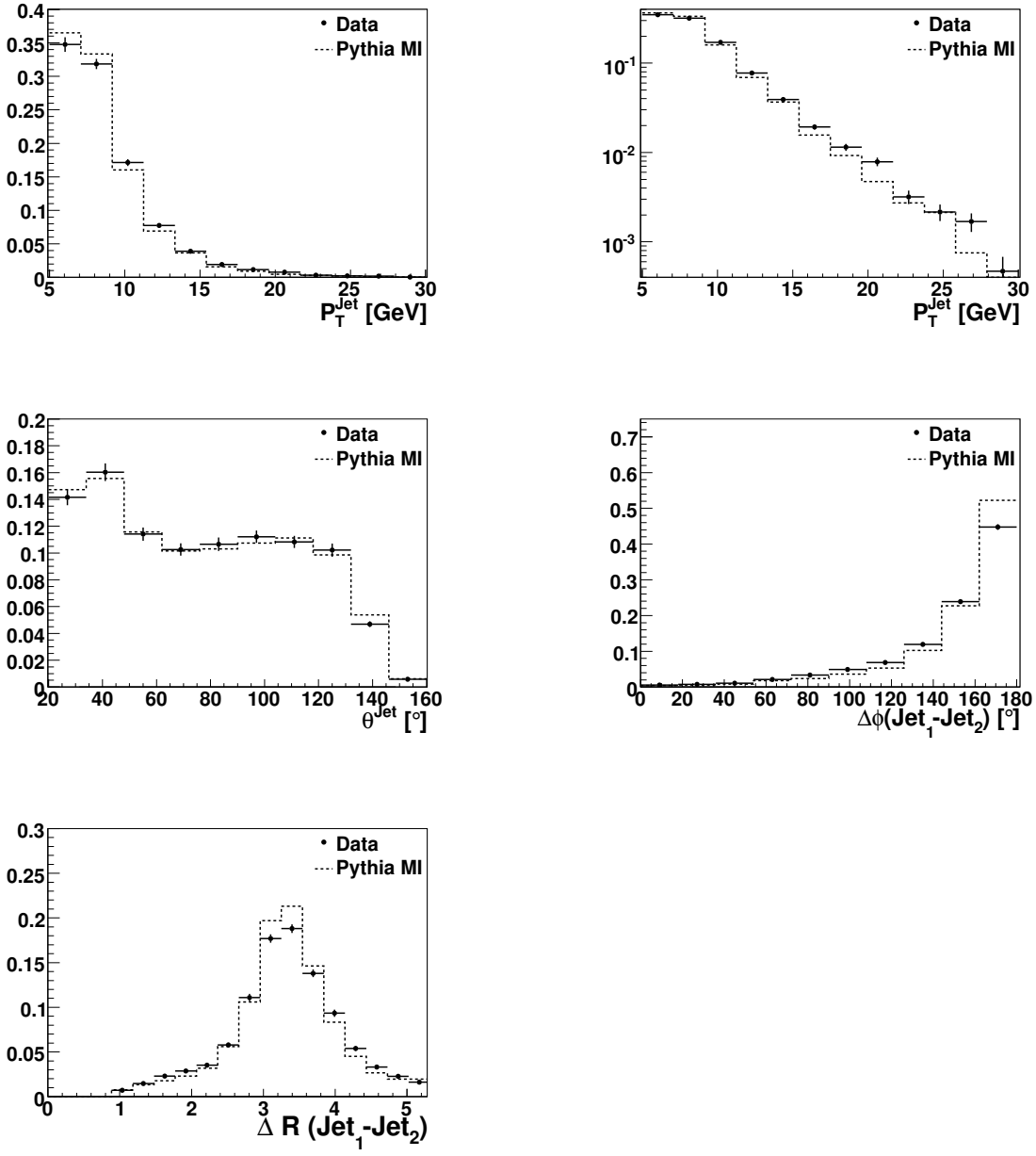


**Figure 4.7:** Leading jets control plots for different variables (see text) after event selection for the charged particle multiplicity studies.

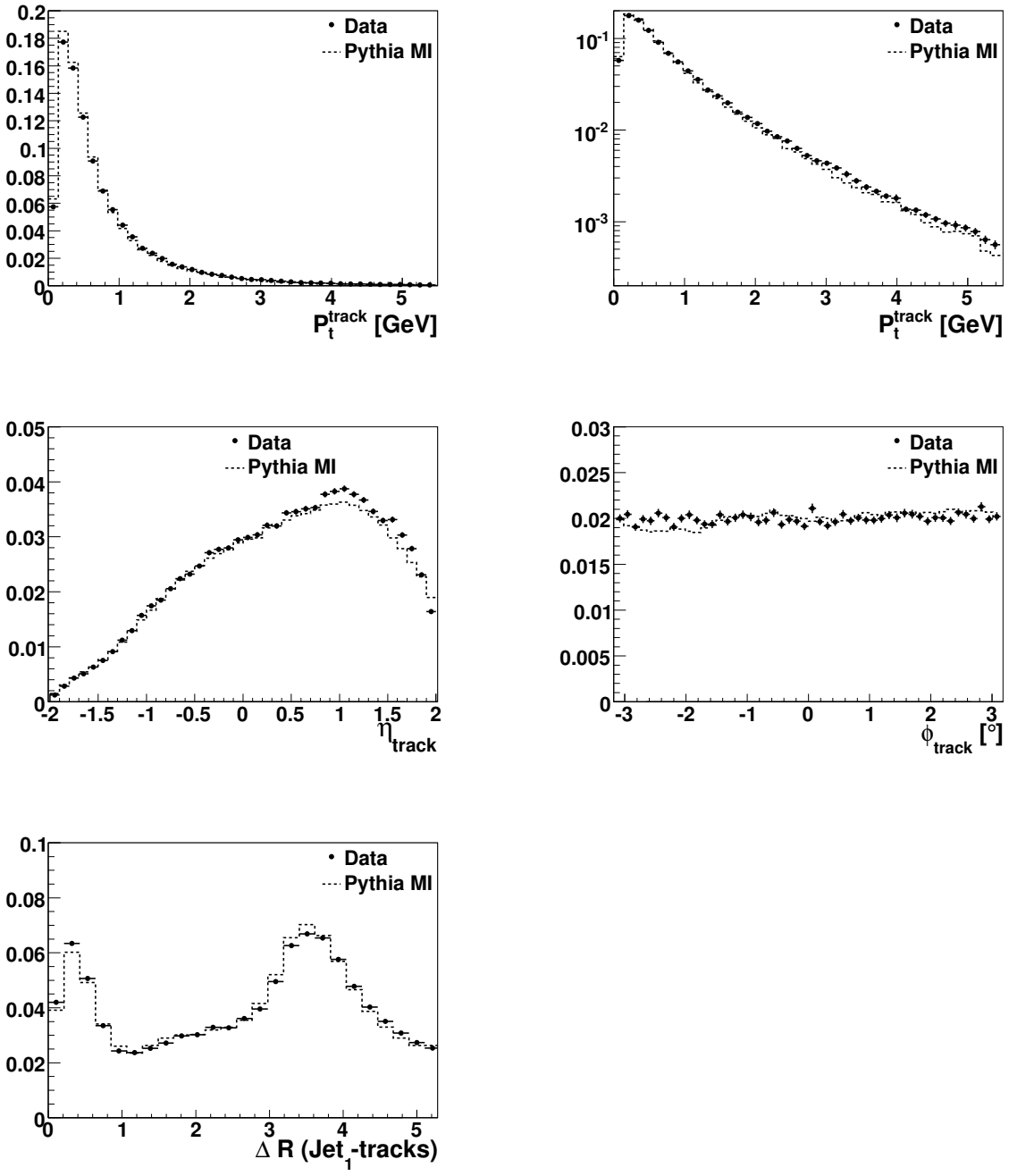


**Figure 4.8:** Control plots for different variables (see text) after event and track selection.

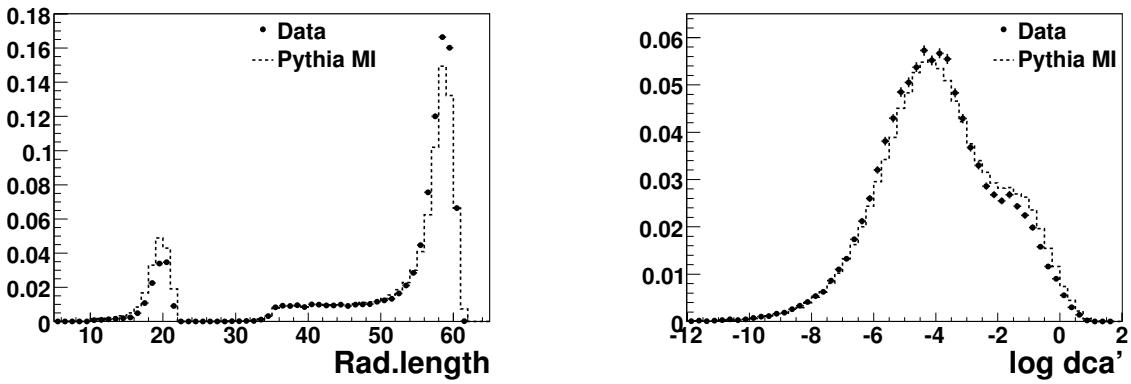


**Figure 4.9:** Control plots for the track length and the logarithm of the actual distance of closest approach distributions are shown after event and track selection.

In figure 4.10 the energy fraction of the photon entering the hard scattering $x_\gamma^{obs}$ (left) and the photon proton centre-of-mass energy $W_{\gamma p}$ (right) distributions are shown. The discrepancies seen in the $x_\gamma^{obs}$ distribution are not important since the detector effects are treated explicitly by the unfolding method (chapter 5).

In figure 4.11 several distributions for the leading jets are shown: the transverse momentum distribution of the leading jet $P_T^{Jet}$ (also in logarithmic scale), the polar angle distribution of the leading jet $\theta^{Jet}$, the azimuthal angle distance of the second leading jet with respect of the leading jet $\Delta\phi(Jet_1 - Jet_2)$ and the $\Delta R(Jet_1 - Jet_2)$.

The discrepancies observed in the transverse momentum of the leading jet, $P_T^{Jet}$, are explicitly treated by the unfolding method. Therefore, they are not expected to affect the data unfolding. The $\Delta\phi(Jet_1 - Jet_2)$ discrepancies were studied when unfolding data by using samples with different $\Delta\phi(Jet_1 - Jet_2)$ distributions. They were also found no to change the results.

In figure 4.12 the transverse momentum spectrum of the minijets in the toward and away regions are shown (upper and lower plots, respectively) for the low and high $x_\gamma^{obs}$ regions (left and right plots, respectively). The data description is reasonable given the statistical uncertainties.

In figure 4.13 the transverse momentum spectrum of the minijets in the transverse high and low activity regions are shown, upper and lower plots, respectively for low and high $x_\gamma^{obs}$ values, left and right plot, respectively. The minijets in data have a larger transverse momentum spectrum than in the MC.

The discrepancies in figures 4.12 and 4.13 are not important if one considers the minijet multiplicity to be the relevant quantity when unfolding and not the transverse momentum of the minijets as done in section 5.9. The minijet multiplicity in the toward and away regions is shown in figure 4.11 (bottom right). The minijet multiplicity in these regions

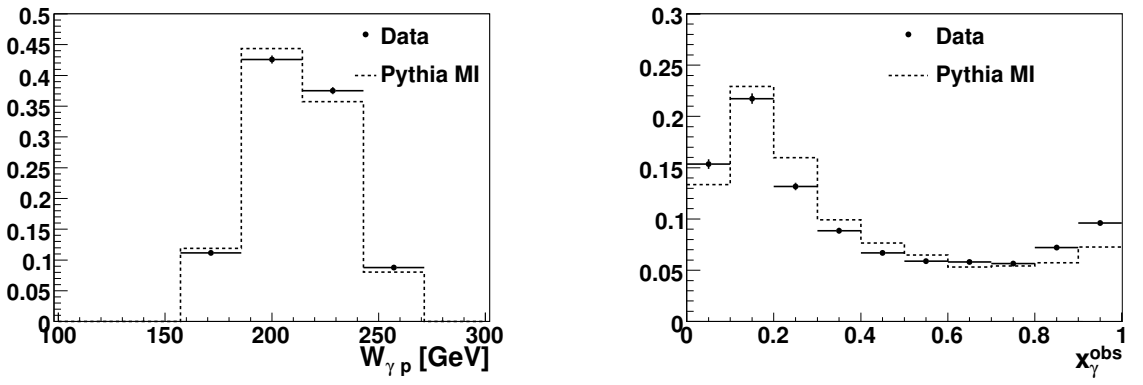


**Figure 4.10:** Energy fraction of the photon entering the hard scattering $x_\gamma^{obs}$ (left) and the photon proton centre-of-mass energy $W_{\gamma p}$ (right) control distributions after event selection for the minijet studies.

is at least one (the leading jet in the toward region) and the data is well described by the MC.

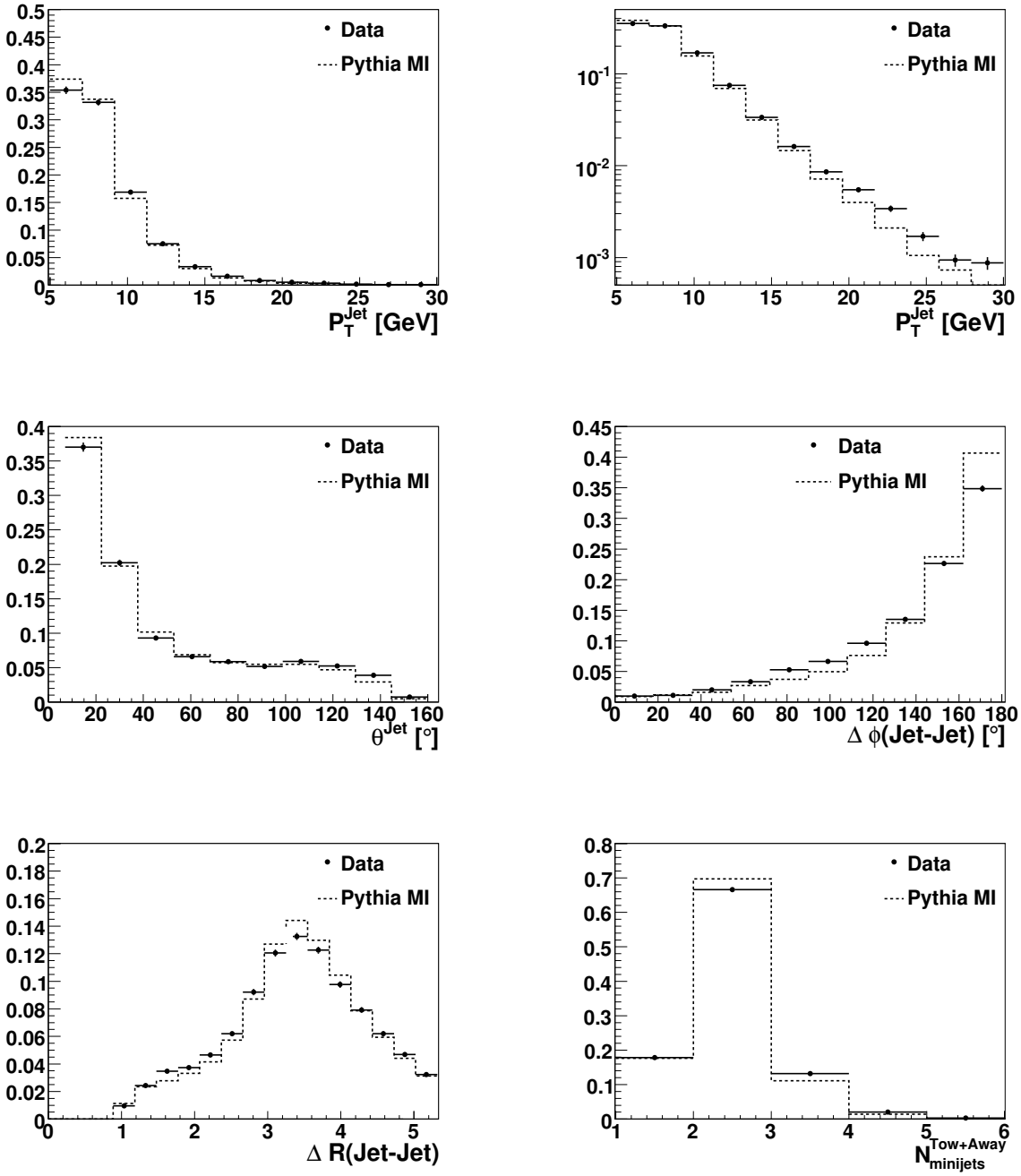


**Figure 4.11:** Leading jets control plots for different variables (see text) after event selection for the minijet multiplicity studies.

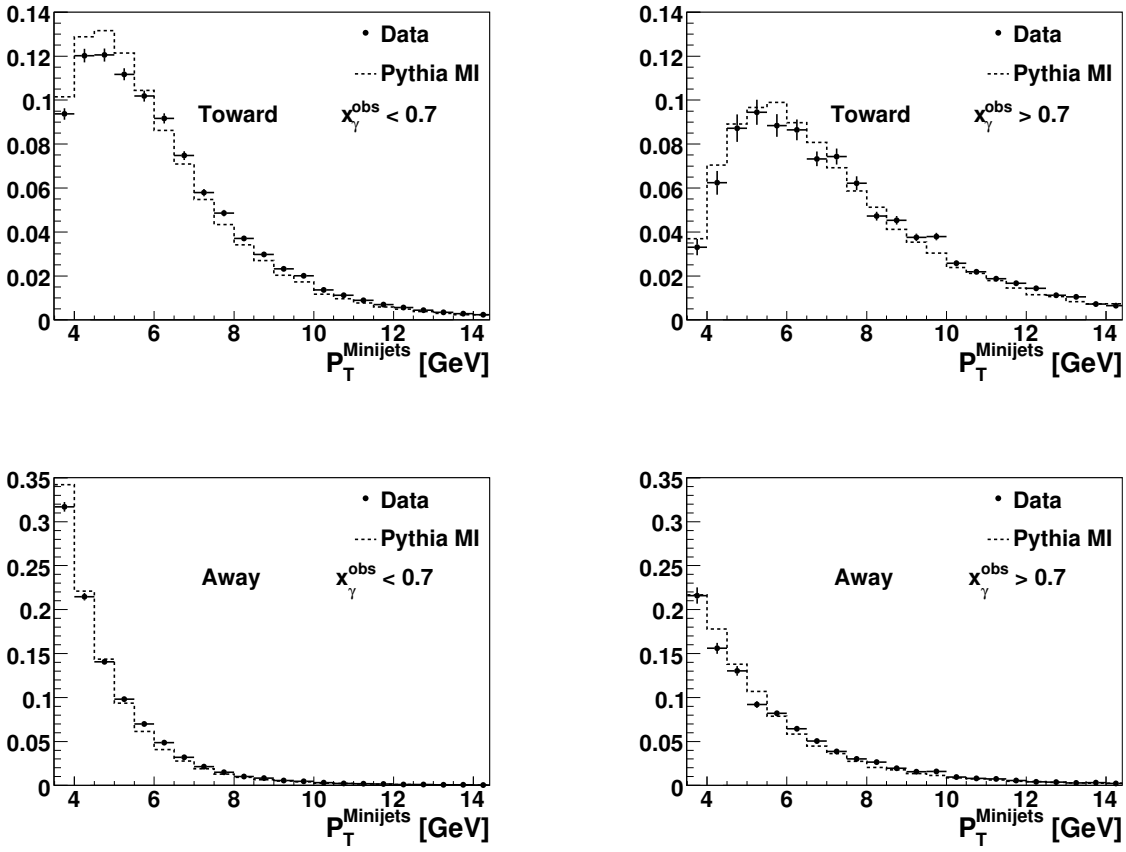


**Figure 4.12:** Transverse momentum spectrum of the minijets in the toward and away regions (upper and lower plots, respectively) for the low and high $x_\gamma^{obs}$ regions (left and right figures, respectively).

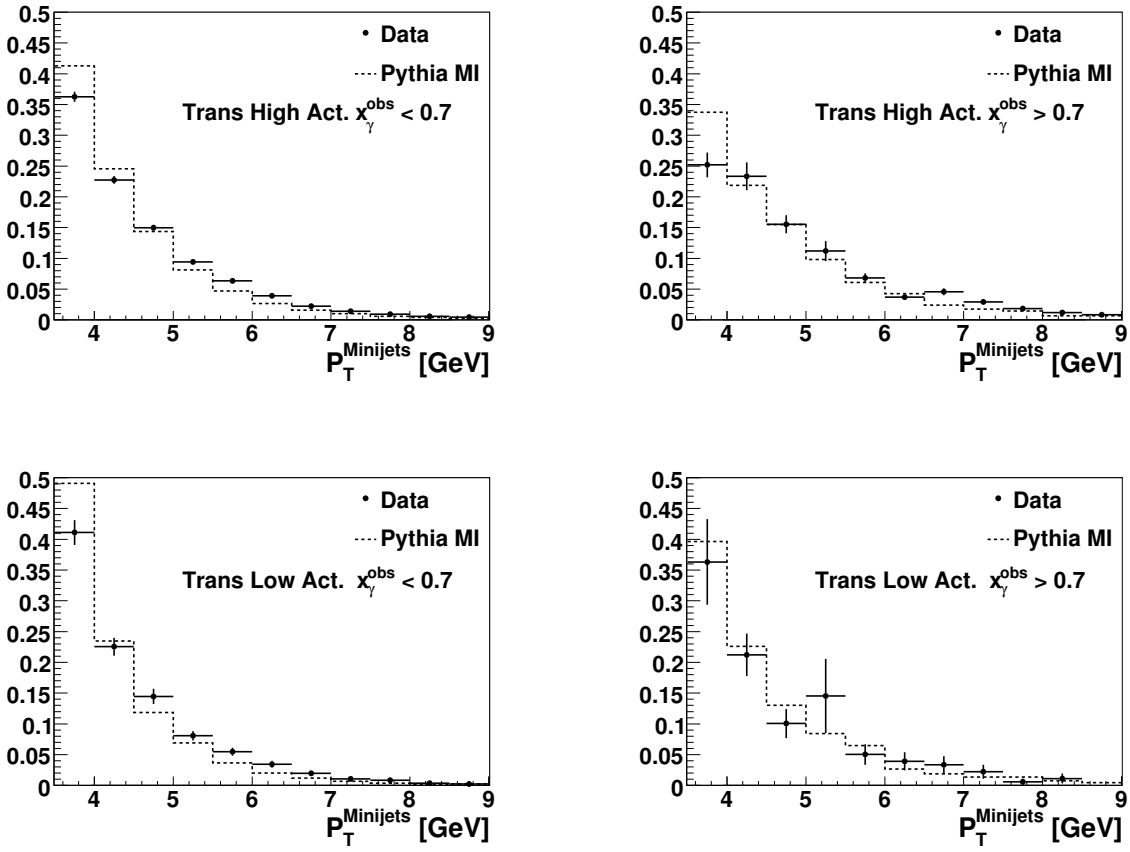


**Figure 4.13:** Transverse momentum spectrum of the minijets in the transverse high and low activity regions (upper and lower plots, respectively) and for the low and high $x_\gamma^{obs}$ regions (left and right figures, respectively).

# Chapter 5

# Unfolding Method

In this chapter we explain unfolding and the importance of unfolding. Two methods for unfolding are presented: the "bin-by-bin" method and a method based on the Bayes' theorem.

## 5.1 Introduction

In a measurement of a physical variable the raw measured values from the detector cannot just be compared to the predictions given by the theory. There are detector related effects and distortions in our measurements which limit the detector performances. This means that the measured distributions are not the "true" distributions and therefore we cannot compare them directly to the theoretical predictions. With a good understanding of our detector we can account for these effects. The goal of unfolding is to get the "true" distributions from the measured data by using all our knowledge on our experimental apparatus.

The "true" distributions correspond to the distributions at hadron level while the distributions from the data directly taken from the detector are the detector level ones.

The unfolding methods rely on detector simulations. They contain all our knowledge about our detector and complement the Monte Carlo event generators, MC, to provide predictions at detector level. Therefore, detector simulations play a key role. From the generated events we have therefore full information about the hadron level, determined from the MC, and the detector level quantities, determined from the detector simulation.

## 5.2 Bin by bin corrections

For a given observable $x$, the "true" distribution, i.e. the distribution at hadron level, $f^{hadr}(x)$, is obtained from the measured ones, $f^{det}(x)$, by using a correction factor $f_{corr}$.

The "true" value in bin $i$ is obtained by:

$$f^{had}(x_i) = f^{det}(x_i) \cdot f_{corr}(x_i) \quad \text{where}$$
$$f_{corr}(x_i) = \left.\frac{f^{had}(x_i)}{f^{det}(x_i)}\right|_{MC}, \tag{5.1}$$

where $f_{corr}$ has to be determined from a MC and a detector simulation. Here we see that this method neglects correlations among bins since the correction factors do not contain explicit information about it. The method is therefore valid only if migrations are negligible. In order to quantify this, it is usual to calculate the *purity* and *stability* of the distributions. In the $i$th bin they are defined as:

$$P^i = \frac{N^i_{had\ \&\ det}}{N^i_{det}} \qquad S^i = \frac{N^i_{had\ \&\ det}}{N^i_{had}}, \tag{5.2}$$

where $N^i_{had}$ and $N^i_{det}$ are the number of events in the $i$th bin of the distribution at hadron and detector level, respectively. $N^i_{had\ \&\ det}$ is the number of events in the $i$th bin present at both levels. Purity is the fraction of events at detector level which are in the same bin as at detector level. Stability is the fraction of events at hadron level which are in the same bin as at hadron level. These definitions help us to see how large the migrations from the other bins or from out of the visible range are in the $i$th bin. Therefore, for the bin-by-bin method to be applicable it is necessary that both purity and stability are large. Otherwise the migrations are too large and the method would not deliver a reliable result.

## 5.3 Bayesian method

This method was developed by G. D'Agostini [65, 66] in order to include the effects of migrations when unfolding and is based on the Bayes' theorem. Other methods are available [67] and deliver similar results as shown for example in [68].

In a given distribution we measure a set of possible effects[1] $E_j\,(j = 1, 2, ..., n_E)$. Each of these effects corresponds to a given bin in our distribution. After $N_{obs}$ number of measurements, we have a histogram of $n(E) = (n(E_1), n(E_2), .., n(E_{n_E}))$ number of measured effects. These are the outcome of a set of causes $C_i\,(i = 1, 2, ..., n_C)$ and $n(C) = (n(C_1), n(C_2), ...n(C_{n_C}))$ are the number of causes. With the help of the Bayes formula, we can express the probability that an effect $E_j$ was produced by a cause $C_i$ as

[1] An effect is what is measured in the detector while a cause is what produces the effect in the detector.

follows:

$$P(C_i|E_j) = \frac{P(E_j|C_i)\ P_0(C_i)}{\sum_{l=1}^{n_C} P(E_j|C_l)\ P_0(C_l)}, \tag{5.3}$$

where $P(E_j|C_i)$, the so called *smearing matrix* **SM**, is the probability that the cause $C_i$ caused an effect $E_j$, $P_0(C_i)$ is the initial probability of the cause $C_i$ to occur and is usually taken from the best knowledge of the physics involved. In $P(E_j|C_i)$ is where our detector knowledge is used. It is evaluated by comparing the hadron level generated according to $P_0(C_i)$ with a simulation of the detector on the hadron level events. The initial number of expected events is:

$$n_0(C_i) = P_0(C_i)N_{obs}. \tag{5.4}$$

The number of events assigned to a cause, $\hat{n}(C_i)$, can be obtained with equation 5.3 :

$$\hat{n}(C_i) = \frac{1}{\epsilon_i}\sum_{j=1}^{n_E} P(C_i|E_j)\ n(E_j), \tag{5.5}$$

where $\epsilon_i$ is the efficiency. If $\epsilon_i = 0$ means that the experiment is not sensitive to our measurement and there is no possibility to use the measurement. The efficiency is expressed as:

$$\epsilon_i = \sum_{l=1}^{n_E} P(E_l|C_i), \tag{5.6}$$

i.e. the sum of all probabilities that a cause $C_i$ on which we are interested will produce an effect $E_j$. Using equations 5.3 and 5.6, equation 5.5 can be rewritten:

$$\begin{aligned} \hat{n}(C_i) &= \frac{1}{\sum_{l=1}^{n_E} P(E_l|C_i)}\sum_{j=1}^{n_E} P(C_i|E_j)\ n(E_j) \equiv \sum_{j=1}^{n_E} M_{ij}\ n(E_j) \\ M_{ij} &\equiv \frac{P(E_j|C_i)P_0(C_i)}{[\sum_{l=1}^{n_E} P(E_l|C_i)]\ \ [\sum_{l=1}^{n_C} P(E_j|C_l)\ P_0(C_l)]} \end{aligned} \tag{5.7}$$

$M_{ij}$ are the elements of the *unfolding matrix* **M**[2]. The number of true total number of

[2]The unfolding matrix is clearly not the inverse of the smearing matrix. Note that all the terms $M_{ij}$ are always positive which would not be the case if **M** would be the inverse of **S**.

events, $\hat{N}_{true}$, is given by:

$$\hat{N}_{true} = \sum_{i=1}^{n_C} \hat{n}(C_i)$$

and the probabilities of the causes to occur can be written as:

$$\hat{P}(C_i) = \frac{\hat{n}(C_i)}{\hat{N}_{true}}. \tag{5.8}$$

One can use a uniform distribution $P_0(C_i) = \frac{1}{n_C}$ and then the detector simulation in order to evaluate the matrix elements of the smearing matrix, $P(E_j|C_i)$, i.e. the probabilities of causes to produce effects. Thus, $\hat{P}(C_i)$ lies between $P_0(C_i)$ and the true one. The closer $P_0$ to the true one is, the closer $\hat{P}(C_i)$ to true will be and suggests an iterative procedure where the number of iterations has to be determined (see also below and section 5.4).

Different MC samples with detector simulation are used to determine the number of iterations needed to unfold the data. The following procedure is performed for each MC sample:

- Choose $P_0$ and calculate $P(E_j|C_i)$, *smearing matrix* with the help of a MC and the detector simulation. $P_0$ is usually the MC hadron level description which describes data at the detector level best.
- From 5.7 and 5.8 calculate $\hat{n}_i(C)$ and $\hat{P}(C)$.
- Calculate $\Delta n^i_{gen}$, where

$$\Delta n^i_{gen} \frac{(\hat{n}_i - n_{gen})^2}{\hat{n}_i + n_{gen}} \tag{5.9}$$

  and $n_{gen}$ is the MC generator level.
- Use $\hat{n}_i(C)$ and $\hat{P}(C)$ as the new values for $n_0(C)$ and $P_0(C)$ in the next iteration.
- If $\Delta n^i_{gen} > \Delta n^{i-1}_{gen}$ stop the iteration procedure. The iteration $i-1$ provides the best results for this sample.

The number of iterations needed to unfold data is determined by the number of iterations needed in the different MC samples, see section 5.4.

Finally, we check that the method converges. We define the factor $\Delta^i_{method}$ as follows:

$$\Delta^i_{method} = \frac{(n_{i-1} - n_i)^2}{n_{i-1} + n_i}. \tag{5.10}$$

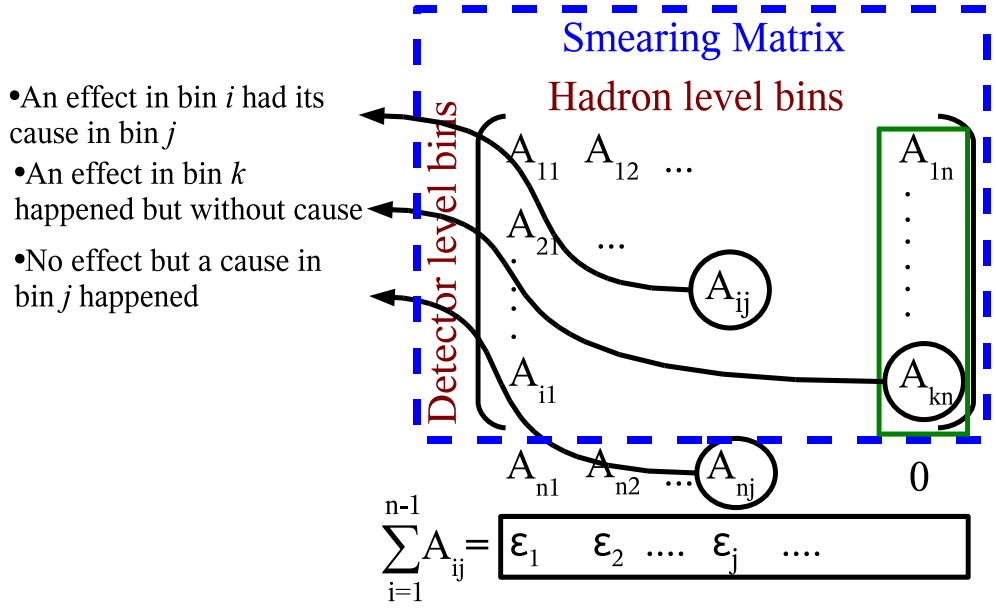


**Figure 5.1:** Each matrix element inside the dashed box, $A_{ij}$, is an element of the smearing matrix $P(E_j|C_i)$, i.e. the probability that a cause $C_i$ caused the effect $E_j$.

It compares the $i$th iteration with the previous one. The method converges when for increasing number of iterations: $\Delta^i_{method} \rightarrow 0$.

## 5.4 Smearing Matrix definitions

In section 5.3 the unfolding method was presented and the smearing matrix (SM) (equation 5.3) and the unfolding matrix (equation 5.7) defined. The SM is very important because it is the basic input to calculate the unfolding matrix. The definition is given in terms of causes, $C_i$, and effects, $E_j$, which correspond to the hadron and the detector level, respectively and have to be specified.

Smearing matrices as the one shown in figure 5.1 are used. Each matrix element inside the dashed box, $A_{ij}$, is an element of the SM and represent the probability $P(E_j|C_i)$ that an effect $E_j$ was measured given the cause $C_i$ occur. The columns and the rows represent detector and hadron level bins, respectively. The elements in the column, $A_{1n} \ldots A_{kn}$, represent the probability that an effect happened but its cause was not explicitly treated in the SM, i.e. the effect has an unknown cause. The sum of probabilities over all the effects produced by a given cause $C_i$ is the efficiency, $\epsilon_i$, equation 5.6. The SM is desired to be diagonal and the elements in the background column small.

In the following sections it will be described how the SM are defined and the causes and effects specified.

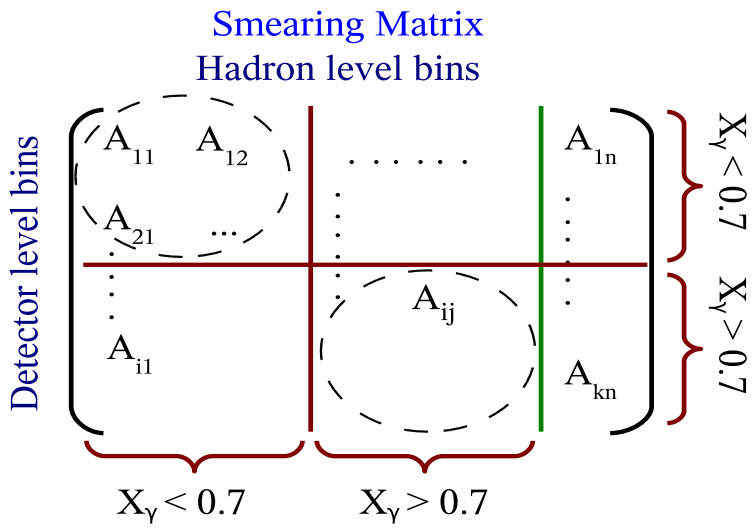


**Figure 5.2:** Correlations between the hadron and detector levels in the $x_\gamma^{obs}$ value are included in the smearing matrix. In an optimal case the elements in the dashed circles should be larger than the others since they represent events with the same $x_\gamma^{obs}$ value at hadron and detector level.

### 5.4.1 Average Number of Tracks Unfolding

The average track multiplicity distribution is unfolded for dijet events with $P_T^{Jets} > 5$ GeV and $-1.5 < \eta^{Jets} < 1.5$. The unfolding is done for two $x_\gamma^{obs}$ regions: $x_\gamma^{obs} < 0.7$ and $x_\gamma^{obs} > 0.7$.

#### Average Track Distributions as a function of $P_T^{Jet_1}$

In order to calculate the average track multiplicity, the number of tracks and number of events are unfolded for each transverse momentum of the leading jet, $P_T^{Jet_1}$, bin and $x_\gamma^{obs}$ region. Once these two quantities are unfolded the average track multiplicity per event is calculated for each $P_T^{Jet_1}$ bin and $x_\gamma^{obs}$ region. This means that two SM are needed.

In order to take into account migrations from low/high $x_\gamma^{obs}$ hadron level values into high/low $x_\gamma^{obs}$ detector level values, smearing matrices which contain information about the $x_\gamma^{obs}$ correlations are used, see figure 5.2. Here, the correlations between the different levels in the $x_\gamma^{obs}$ value are included in the SM. In an optimal case, the elements in the dashed circles and especially the diagonal elements, should be larger than the others since they represent events with the same $x_\gamma^{obs}$ value at hadron and detector level.

Since the $P_T^{Jet_1}$ distribution grows for decreasing $P_T^{Jet_1}$ (figure 4.7), migrations from events with $P_T^{Jets} < 5$ GeV could be important when unfolding and are taken explicitly into consideration. This is done by including in the SM effects and causes were both

leading jets fulfil $P_T^{Jets} > 3.5$ GeV. This means that, when unfolding, the jet selection in table 4.3 is modified to dijet events with $P_T^{Jets} > 3.5$ GeV and from now on in this chapter event selection will include this modified jet selection. This information is included within the brackets in figure 5.2, see also below.

To **unfold the number of events** the elements of the SM are evaluated as follows:

- The element $A_{11}$ is the probability that an event is in the range $x_\gamma^{obs} < 0.7$ and same $P_T^{Jet_1}$ bin, in fact $3.5 < P_T^{Jet_1} < 5$ GeV, at detector and hadron level.
- The element $A_{12}$ ($A_{21}$) is the probability that an event is in the range $x_\gamma^{obs} < 0.7$ at detector and hadron level but different $P_T^{Jet_1}$ bin: at hadron (detector) level $5 < P_T^{Jet_1} < 8.5$ GeV while at detector (hadron) level $3.5 < P_T^{Jet_1} < 5$ GeV.
- The element $A_{i1}$ is the probability that an event is in a different $x_\gamma^{obs} < 0.7$ range and $P_T^{Jet_1}$ bin at detector and hadron level.
- The element $A_{1n}$ is the probability that an event at detector level is in the region $x_\gamma^{obs} < 0.7$ and $3.5 < P_T^{Jet_1} < 5$ GeV but at hadron level the event was not selected.
- If an event was selected at hadron level but not a detector level this is taken into account by the efficiencies $\epsilon_i$ in figure 5.1 (equation 5.6).

An event not present neither at detector nor at hadron level does not need to be taken into account in the SM because it does not imply any migration or detector effect to be corrected.

The unfolding matrix obtained is shown in figure 5.3. In this unfolding matrix the correlations between different $x_\gamma^{obs}$ regions at hadron and detector level and the background from dijet events with $P_T^{Jets} > 3.5$ GeV are taken explicitly into account.

To **unfold the number of tracks** a more complex SM is needed because the different regions in the transverse plane, i.e. the toward, away and the two transverse regions (section 2.4), have to be considered.

For simplicity we describe how this is done using two regions only. In figure 5.4 the SM for two of these regions are shown, 'region 1' and 'region 2'. The generalization to four regions is straight forward. The big dashed box in figure 5.4 is divided into four boxes: two solid boxes of medium size in the diagonal and two dashed boxes of the same size out the diagonal. These four boxes are in turn divided into four small boxes: two solid boxes in the diagonal and two dashed boxes out of the diagonal. Finally, for events selected only at detector level there is a background column (on the right, solid rectangle).

The procedure to evaluate the SM is done as follows:

A If an event was selected at both levels consider tracks at detector level and charged particles at hadron level as follows:

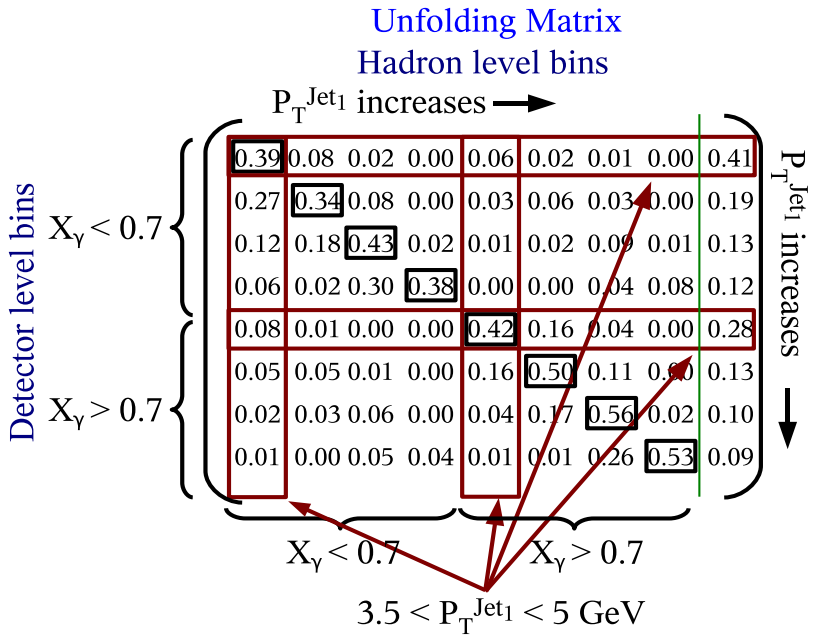


**Figure 5.3:** Example of unfolding matrix for the number of events. The correlations between different $x_\gamma^{obs}$ regions at hadron and detector level and the background from dijet events with $P_T^{Jets} > 3.5$ GeV are taken explicitly into account.

A1 If a charged particle and its correspondent track exist and passed the selection at their corresponding levels they are paired and contribute to the big dashed box. Different possibilities exist:

- If the event was classified to be in the same $x_\gamma^{obs}$ region at both levels, the track-charged particle pair contribute to the medium solid boxes. Two further possibilities exist:
  - I If the track belongs to the same transverse plane region, here 'region 1' or 'region 2', than the charged particle: the track-charged particle pair contribute to the **small solid boxes within the two medium solid boxes**.
  - II If the track belongs to a different transverse plane region than the charged particle: the track-charged particle pair contribute to the **small dashed boxes within the two medium solid boxes**.
- If the $x_\gamma^{obs}$ region is different for the detector and the hadron levels the track-charged particle pair contribute to the medium dashed boxes.
  - I If the track belongs to the same transverse plane region than the charged particle: the track-charged particle pair contribute to the **small solid boxes within the two medium dashed boxes**.

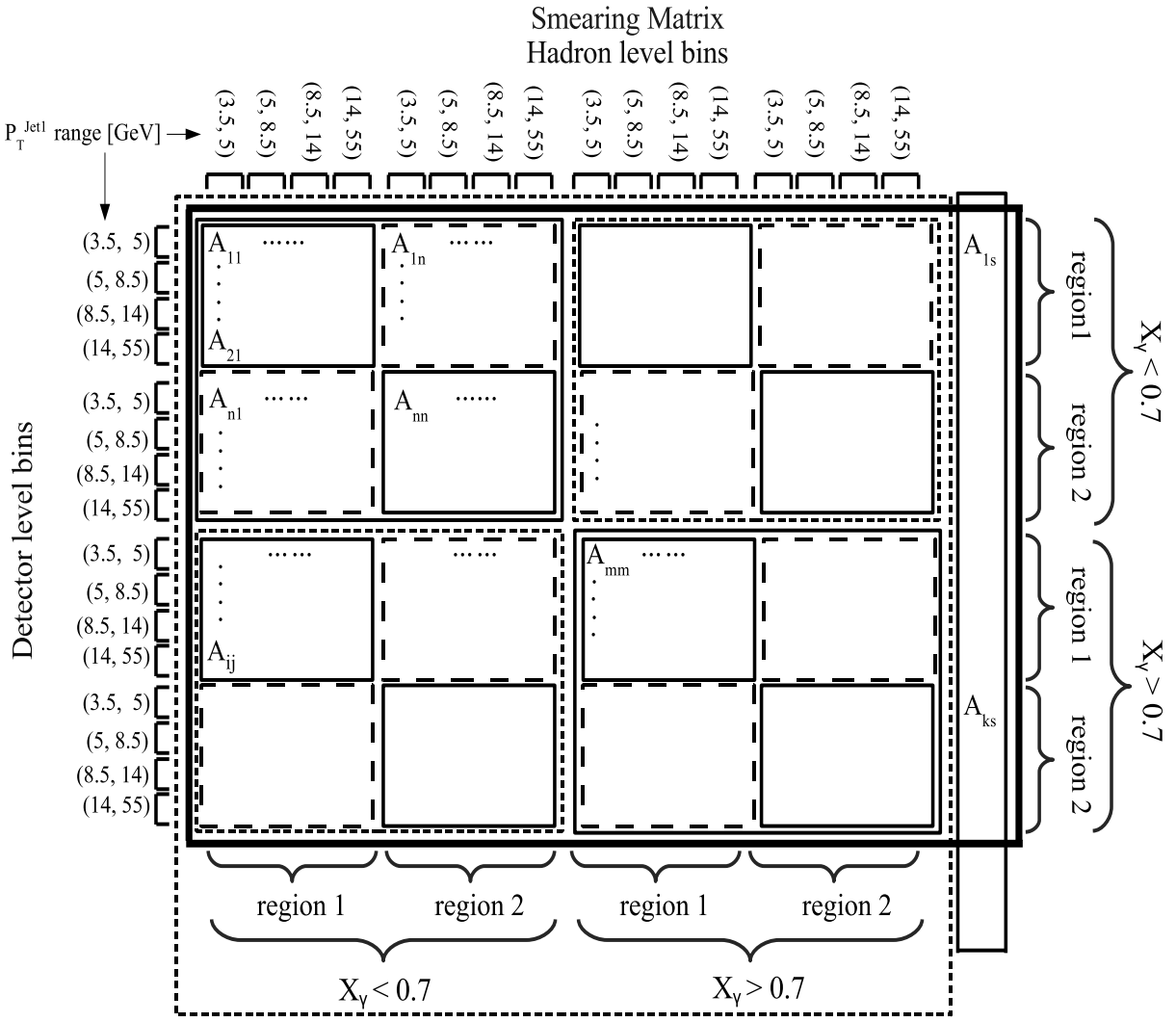


**Figure 5.4:** The figure shows how the SM to unfold the number of tracks is organized. The big dashed box is divided into four boxes of medium size: two solid boxes in the diagonal and two dashed boxes out the diagonal. These four boxes are in turn divided into four small boxes: two solid boxes in the diagonal and two dashed boxes out of the diagonal (see explanations in text).

II If the track belongs to a different transverse plane region than the charged particle: the track-charged particle pair contribute to the **small dashed boxes within the two medium dashed boxes**.

A2 If a track passed the selection but the charged particle did not or does not exist the track contribute to the background column in the **solid rectangle**.

A3 If the charged particle passed the selection but not the track or the charged particle did not produce any track this contributes to the **efficiencies**, $\epsilon_i$.

A4 When all tracks and charged particles were considered the next event is considered.

C If an event was selected only at detector level all tracks in the event contribute to the background column in the **solid rectangle**. The next event is considered.

D If an event was selected only at hadron level all the charged particles contribute to the **efficiencies**, $\epsilon_i$. The next event is considered.

The obtained unfolding matrix is shown in figure 5.5.

Migrations from high/low $x_\gamma^{obs}$ hadron level regions to low/high $x_\gamma^{obs}$ regions at detector level and migrations from different $P_T^{Jet_1}$ bins are treated explicitly in the SM but other migrations are also included.

If the leading jet at hadron level is not at the same azimuthal angle than the leading jet at detector level the event looks rotated. Figure 5.6 shows the difference in the azimuthal angle of the leading jet at hadron and detector level. The figure shows that the rotation is usually less than 0.2 radians (central peak) but in many events the rotation is of about $\pi$, i.e. the leading jet at one level is the second hardest jet at the other level. The effect of such a rotation is that the toward and away regions are exchanged from one level to the other. Since the rotation is not exactly 180° the rotation may also affect the definition of the transverse regions. Since all regions are included in one single SM, migrations toward-away and high-low transverse region as well well as any other combination are taken into account.

### Unfolding Tests: Smearing Matrix elements statistical precision and number of iterations.

Two tests on unfolding are presented: SM elements statistical precision and number of iterations determination. The average track distributions as a function of $\Delta\phi^{Jet_1,\ h^\pm}$ in two $x_\gamma^{obs}$ regions exemplify the examples.

It is important to check that the elements of the SM are well known, i.e the elements are determined with a good statistical precision. A hadron level prediction was used where the detector effects in the events were simulated, a MC simulation. This sample was then divided in several smaller samples (usually of the same size than the data sample to be

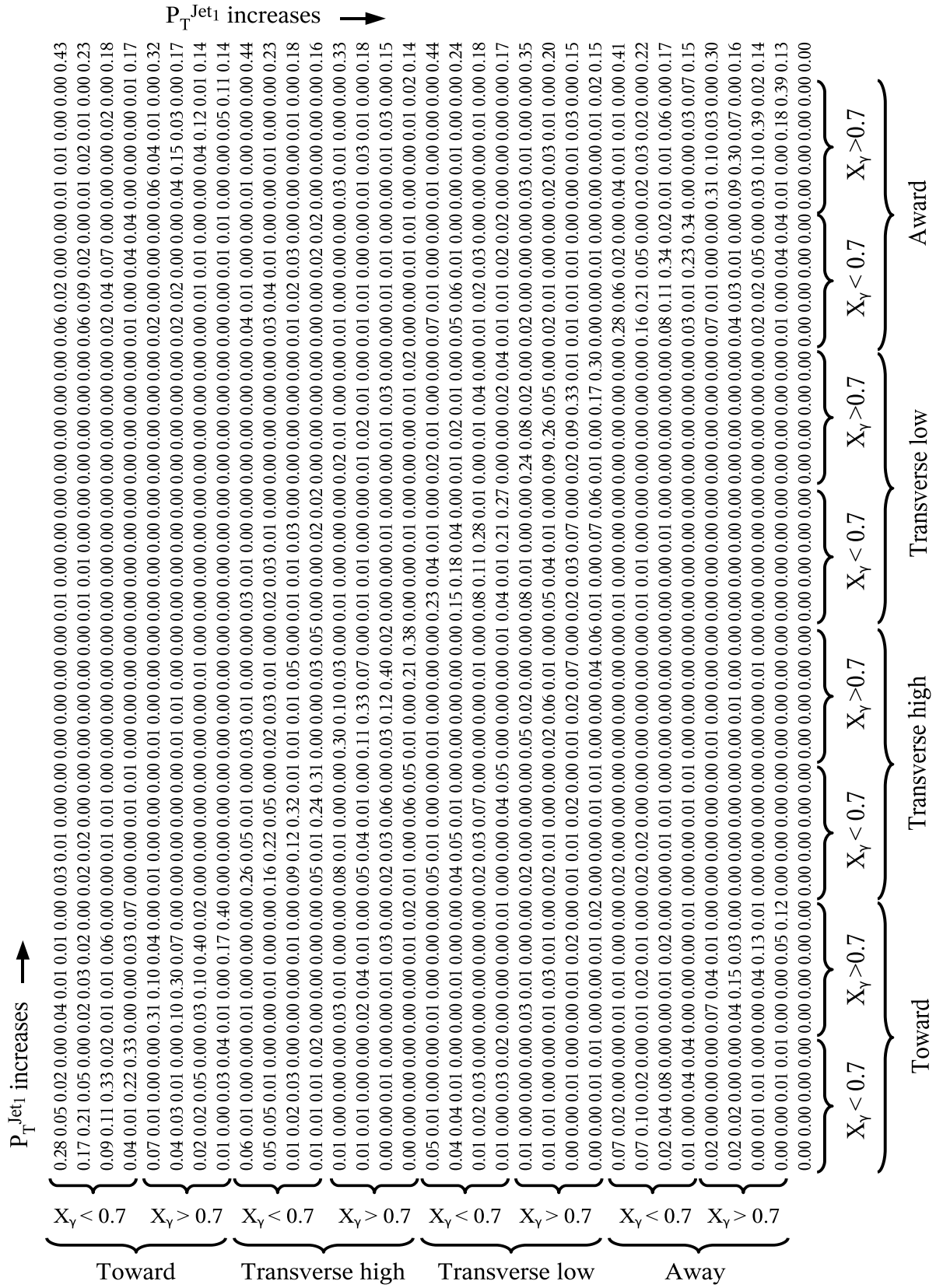


**Figure 5.5:** Unfolding matrix for the number of tracks.

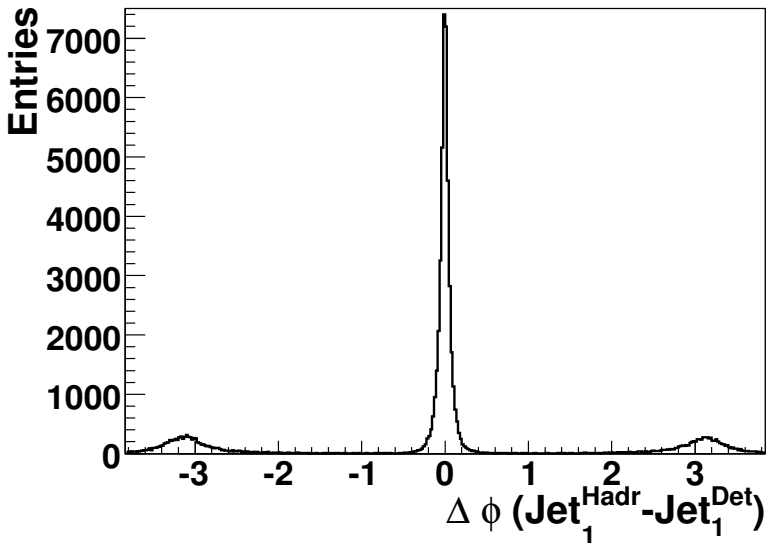


**Figure 5.6:** Difference in the azimuthal angle of the leading jet at hadron and detector level.

unfolded). Each of these subsamples were used to unfold the others. If it is is found that the SM is well determined with a subsample then it means that entire MC sample has enough statistics for the final SM calculation.

In figure 5.7 a MC sample (MC-1) was used to unfold another sample with similar statistics (both with similar statistics than the data). The unfolded track multiplicity is shown in the two $x_\gamma^{obs}$ regions: $x_\gamma^{obs} < 0.7$ (left) and $x_\gamma^{obs} > 0.7$ (right). The leading jet axis is situated at $\Delta\phi^{Jet_1,h\pm} = 180°$ (section 2.4). It is shown that the unfolded distribution (uf) is very close to the MC generator level, i.e the hadron level. This is especially visible in the $x_\gamma^{obs} < 0.7$ region.

The Bayesian method for unfolding is iterative and the number of iterations needed has to be determined using different MC samples. To check this, the MC samples have to be different statistically, as in the example above, and different at hadron level. It is checked that a sample (MC-2) with a different hadron level[3]than MC-1 can be unfolded with the same number of iterations[4]. The result is shown in figure 5.8. The MC-2 charged particle multiplicity values at $\Delta\phi^{Jet_1,\ h\pm} \approx 180°$ and $\Delta\phi^{Jet_1,\ h\pm} \approx 0°$ are higher than the MC-1 (up to 10%). Regardless the rather large differences, the unfolded distribution (MC-2 Uf) is very close to the generator level (MC-2 Gen).

Other test were also done. Weights on variables like $P_T^{Jet_1}$ or $\eta^{Jets}$ were used in the samples

[3]The event selection for MC-2 was $P_T^{Jet_1} > 7$ GeV and $P_T^{Jet_1} > 6$ GeV.

[4]In case they differ it can be taken as a systematic source of uncertainty. Since the method converges, section 5.3, it is reasonable guess.

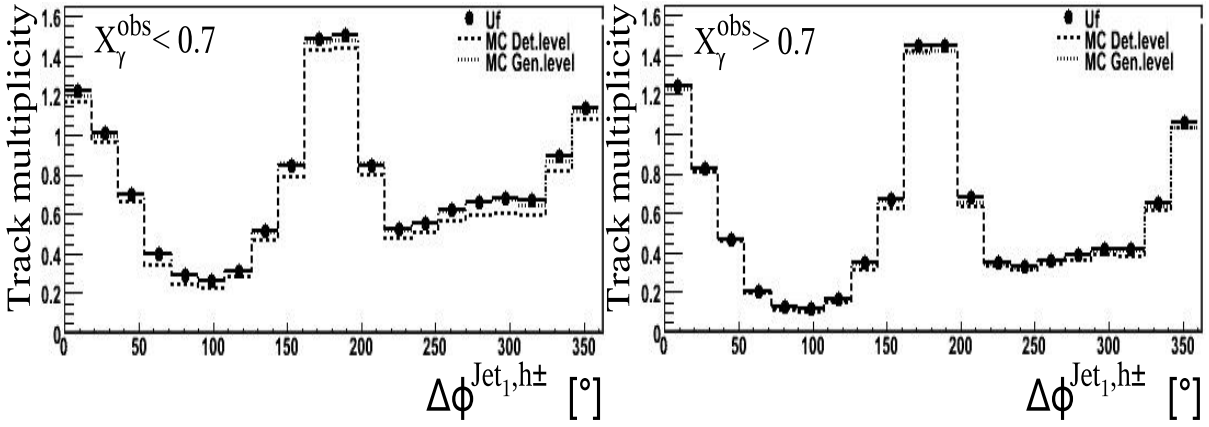


**Figure 5.7:** A MC sample (MC-1) was used to unfold another sample with similar statistics. The unfolded distribution (uf) is very close to the MC generator level. The first iteration gave the best results.

and then unfolded. The first iteration gave in all cases the best results.

### 5.4.2 Average Number of Minijets Unfolding

The average minijet multiplicity distribution is unfolded for dijet events with $P_T^{Jets} > 5$ GeV and $-1.5 < \eta^{Jets} < 2.79$. The unfolding is done in two $x_\gamma^{obs}$ regions: $x_\gamma^{obs} < 0.7$ and $x_\gamma^{obs} > 0.7$. For the minijets only the average multiplicity distribution as a function of $P_T^{Jet_1}$ is unfolded.

At hadron level jets are correlated to partons but the definition of jet is not unique (section 1.7). Similarly, jets at detector level are correlated to jets at hadron level but it is rather ambiguous to identify hadron with detector level jets. For this reason, it is not possible to define causes and effects as it was done for tracks. The minijet multiplicity is compared at detector level with the hadron level in each transverse plane region, i.e. no correlations between transverse plane regions are considered and four independent SM are used.

The procedure to fill one of the transverse plane regions SM (figure 5.9) is the following:

- If an event was selected at hadron as well as at detector level and are in the same $x_\gamma^{obs}$ region, this event contributes to the solid boxes, i.e. the upper left solid big box and bottom right solid big box for low and high $x_\gamma^{obs}$ values, respectively.

  If at both levels the $P_T^{Jet_1}$ is in the same bin, the event contributes to the small solid boxes along the diagonal of the SM. If the minijet multiplicity is the same the

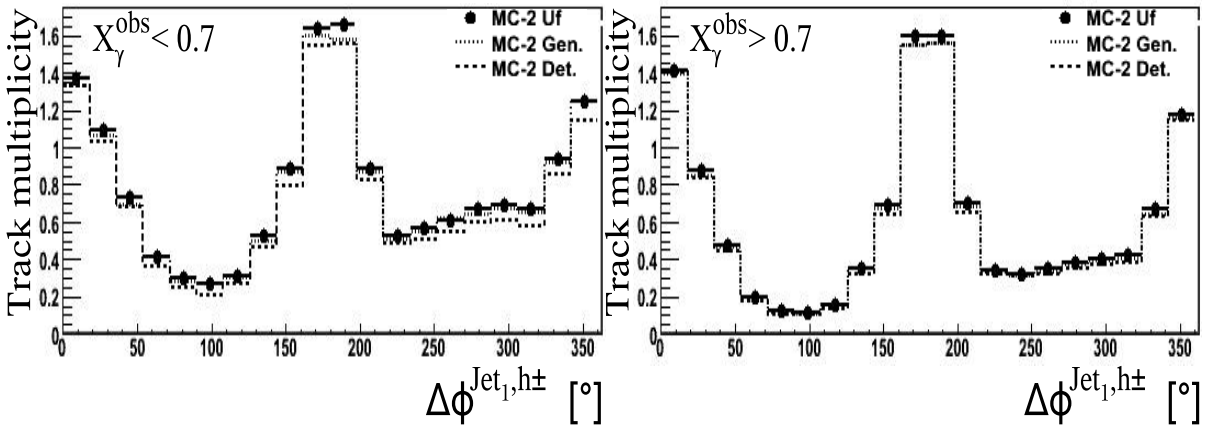


**Figure 5.8:** A MC sample (MC-1) was used to unfold another sample (MC-2) which has a different hadron level. The unfolded distribution (MC-Uf) is very close to the MC generator level (MC-2 Gen). In this case the first iteration also gave the best results.

event contributes to the diagonal elements, for example $A_{11}$, while if not the event contributes to the off diagonal elements inside the solid box, for example $A_{32}$.

In the toward regions there is by definition always at least one jet with transverse momentum larger than 3.5 GeV and therefore the matrix elements $A_{1i}$ and $A_{i1}$ are zero, i.e. probability zero.

- If an event was selected at hadron as well as at detector level but they are in different $x_\gamma^{obs}$ regions, this event contributes to the dashed boxes.
- If an event was selected at detector level but not at hadron level this event contributes to the solid rectangle background column.
- If an event was selected at hadron level but not at detector level this event contributes to the efficiencies, $\epsilon_i$.

Once the events are unfolded, the average minijet multiplicity is calculated as follows in each transverse plane region:

$$\langle N \rangle = \frac{\sum_{i=0}^{4} i\ n_i}{\sum_{i=0}^{4} n_i}, \tag{5.11}$$

where the sum runs over the possible minijet multiplicity, and $n_i$ is the number of the unfolded events in a given transverse plane and $x_\gamma^{obs}$ regions and $P_T^{Jet_1}$ bin.

The optimal number of iterations was found to be two iterations in all cases.

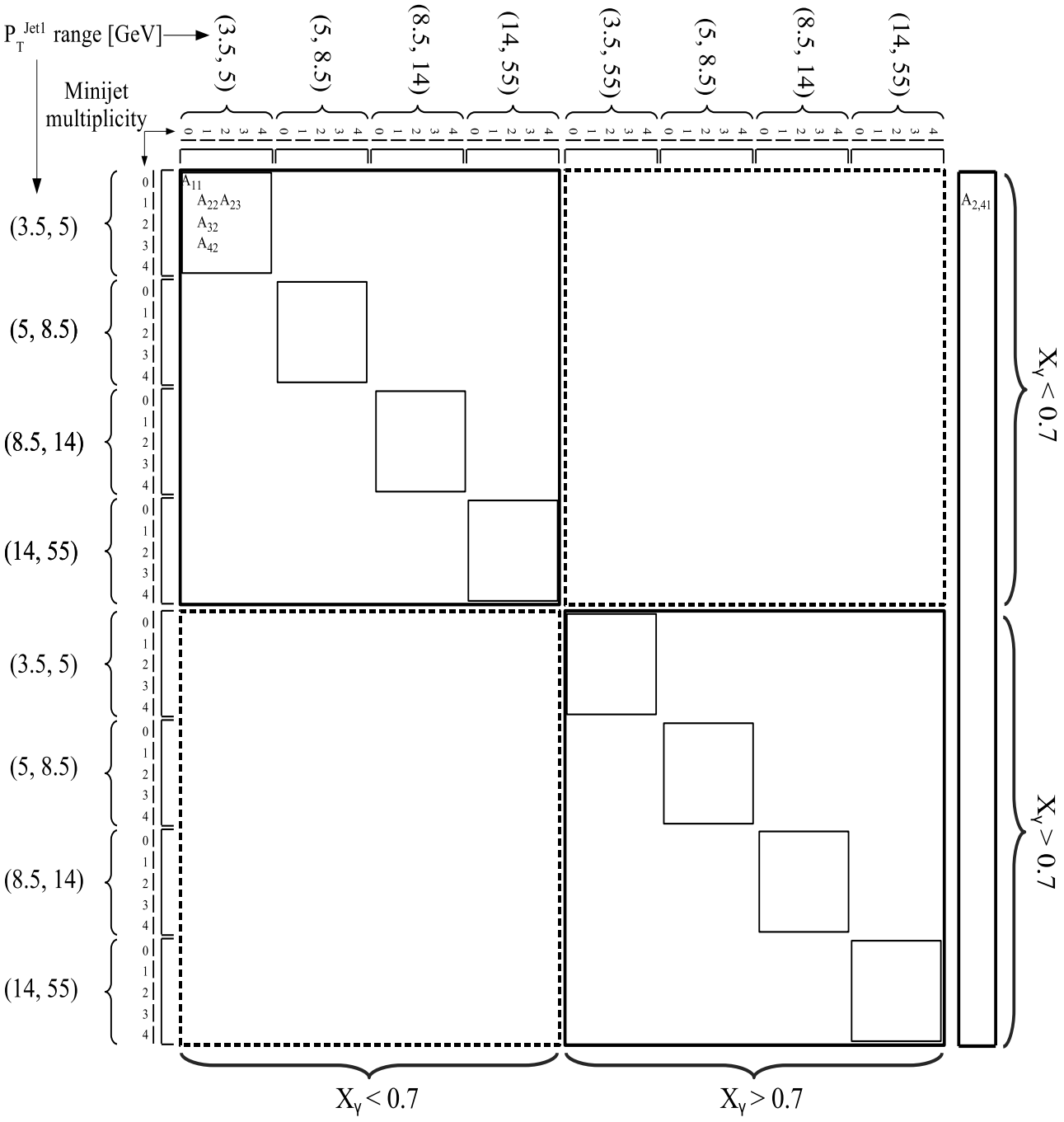


**Figure 5.9:** A MC sample (MC-1) was used to unfold another sample (MC-2) which has a different hadron level. The unfolded distribution (MC-Uf) is very close to the MC generator level (MC-2 Gen). In this case the first iteration also gave the best results.

# Chapter 6

# Average Charged Particle and Minijet Multiplicities

In this chapter the measurement of the observables presented in section 2.4, i.e. average track and minijet multiplicities are presented at hadron level.

## 6.1 Measurement Uncertainties

The unfolding method used (section 5.3) provides two sources of uncertainties [65,66]: the statistical uncertainty due to the finite number of measured effects and the uncertainty on the elements of the smearing matrix, see equation 5.7. Another source of uncertainty could be present, namely the initial guess of the probabilities of the causes $P_0(C_i)$. Since both for the track and the minijet unfolding a fixed number of iterations was found, no uncertainty is associated to $P_0(C_i)$.

In addition, systematic uncertainties are present. These uncertainties are added in quadrature with the uncertainties provided by the unfolding method.

#### Hadronic energy scale uncertainty

The uncertainty in the hadronic energy scale was found to be less than $\pm 2\%$ (section 3.4). The hadronic final state energy was varied in the data by this amount both for the track and the minijet studies.

#### Tracking uncertainty

The following sources of uncertainty for the tracks were considered: $\pm 2\%$ for the track finding in the CJC, $\pm 1\%$ for the track vertexing in the CJC and $\pm 1.5\%$ for the nuclear

interaction simulation. Adding all contributions in quadrature, the total tracking uncertainty is found to be ±2.7% and this is applied as an independent and uncorrelated uncertainty over the $P_T$ and $\eta$ range of the tracks [69].

**Transverse High/Low activity region resolution**

The definition of high/low activity region could depend on whether a track was missing. The uncertainty due to this fact was calculated by assigning a 10% probability of misidentification of the high and low activity regions when the difference in the $P_T^{sum}$ (section 2.4) is less than 160 MeV.

In the case of the minijet measurement, the uncertainty of high/low activity region misidentification is equivalent to the hadronic energy scale uncertainty and therefore no specific uncertainty is taken into account.

## 6.2 Charged Particle Multiplicity

In this section, the measurement of average charged particles with $P_T^{ch} > 150$ MeV and $|\eta^{ch}| < 1.5$ in dijet events with $P_T^{Jets} > 5$ GeV and $|\eta^{Jets}| < 1.5$ in the photoproduction regime $Q^2 < 0.01$ GeV$^2$ is presented. The measured values and their uncertainties can be found in appendix B. The parameters in PYTHIA can be found in appendix D.

### 6.2.1 Average track multiplicity as a function of $\Delta\phi^{Jet_1,\ h\pm}$

The average track multiplicity as a function of the azimuthal angle with respect the leading jet, $\Delta\phi^{Jet_1,\ h\pm}$, is shown in figure 6.1 in two $x_\gamma^{obs}$ regions: a resolved photon enriched region with $x_\gamma^{obs} < 0.7$ (left) and a photon point-like enriched region with $x_\gamma^{obs} > 0.7$ (right).

The leading jet is set to be at $\Delta\phi^{Jet_1,\ h\pm} = 180°$, centered in the toward region $120° < \Delta\phi^{Jet_1,\ h\pm} < 240°$. Similarly defined are the transverse high activity region in the range $240° < \Delta\phi^{Jet_1,\ h\pm} < 300°$, the low activity region in the range $60° < \Delta\phi^{Jet_1,\ h\pm} < 120°$ and the away region in $300° < \Delta\phi^{Jet_1,\ h\pm} < 60°$. The second hardest jet is usually situated in the away region t. If in a given event the transverse high and low activity region are not in these ranges the event is rotated.

The data are compared to PYTHIA with and without multiple parton interactions (PYTHIA MI and PYTHIA NMI, respectively). In PYTHIA, the DGLAP parton evolution equations and the Lund string model fragmentation are implemented. The largest average charged particle multiplicity corresponds to the toward region, where the leading jet is located. In this region, the average charged particle multiplicity is similar in both $x_\gamma^{obs}$ regions. Between the transverse regions, the high activity regions have larger average charged particle multiplicities than those in the low activity regions.

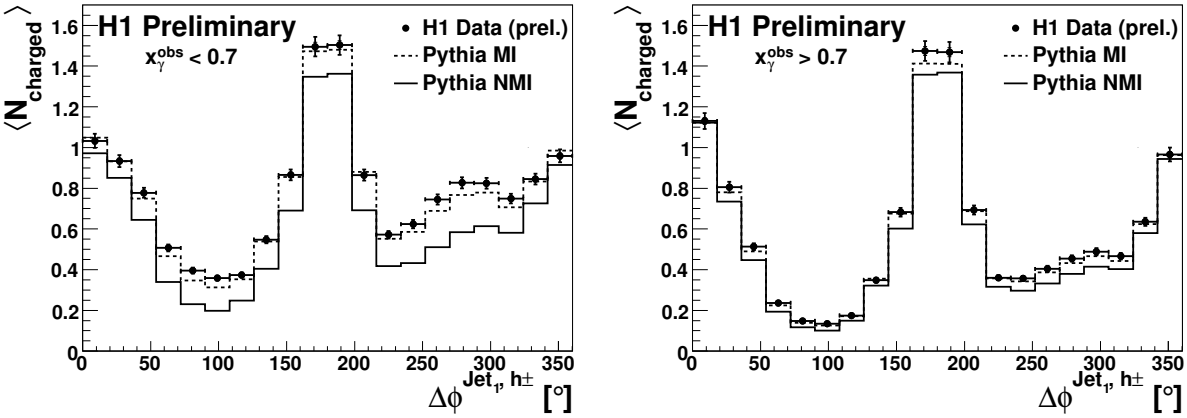


**Figure 6.1:** Average track multiplicity as a function of the azimuthal angle with respect the leading jet, $\Delta\phi^{Jet_1,\, h^\pm}$. The leading jet is set to be at $\Delta\phi^{Jet_1,\, h^\pm} = 180°$ and the transverse high activity region in the range $240° < \Delta\phi^{Jet_1,\, h^\pm} < 300°$. Two $x_\gamma^{obs}$ regions are shown: a resolved photon enriched region with $x_\gamma^{obs} < 0.7$ (left) and a photon point-like enriched region with $x_\gamma^{obs} > 0.7$ (right). Data are compared to PYTHIA with and without the MPI model.

The average charged particle multiplicity description given by PYTHIA MI is larger than PYTHIA NMI, especially at low $x_\gamma^{obs}$ values (left), where the contribution from resolved photon events is large (section 1.4). In this $x_\gamma^{obs}$ region, the inclusion of multiple parton interactions (MPI) which contributes as a pedestal, improves the description of the data.

In the high $x_\gamma^{obs}$ region (right), the differences between both descriptions are small. The differences come from resolved events still present at $x_\gamma^{obs} > 0.7$ values and PYTHIA MI gives also here a better description of the data.

### 6.2.2 Average track multiplicity as a function of $P_T^{Jet_1}$

The average charged particle multiplicity as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$, for the toward and away regions (upper and lower plots, respectively) and for the low and high $x_\gamma^{obs}$ regions (left and right, respectively) is shown in figure 6.2. The average charged particle multiplicity rises with increasing $P_T^{Jet_1}$ from around 4-5 charged particles at $P_T^{Jet_1} \sim 5$ GeV to around 7 charged particles at $P_T^{Jet_1} \sim 15$ GeV also depending on $x_\gamma^{obs}$. In the region $x_\gamma^{obs} > 0.7$, the measurements are reasonably well described by PYTHIA NMI, whereas at $x_\gamma^{obs} < 0.7$ this is clearly not enough, especially at the lower values of $P_T^{Jet_1}$. Including MPI brings the prediction in good agreement with the measurement.

In figure 6.3 the average charged multiplicity in the transverse regions is shown for the

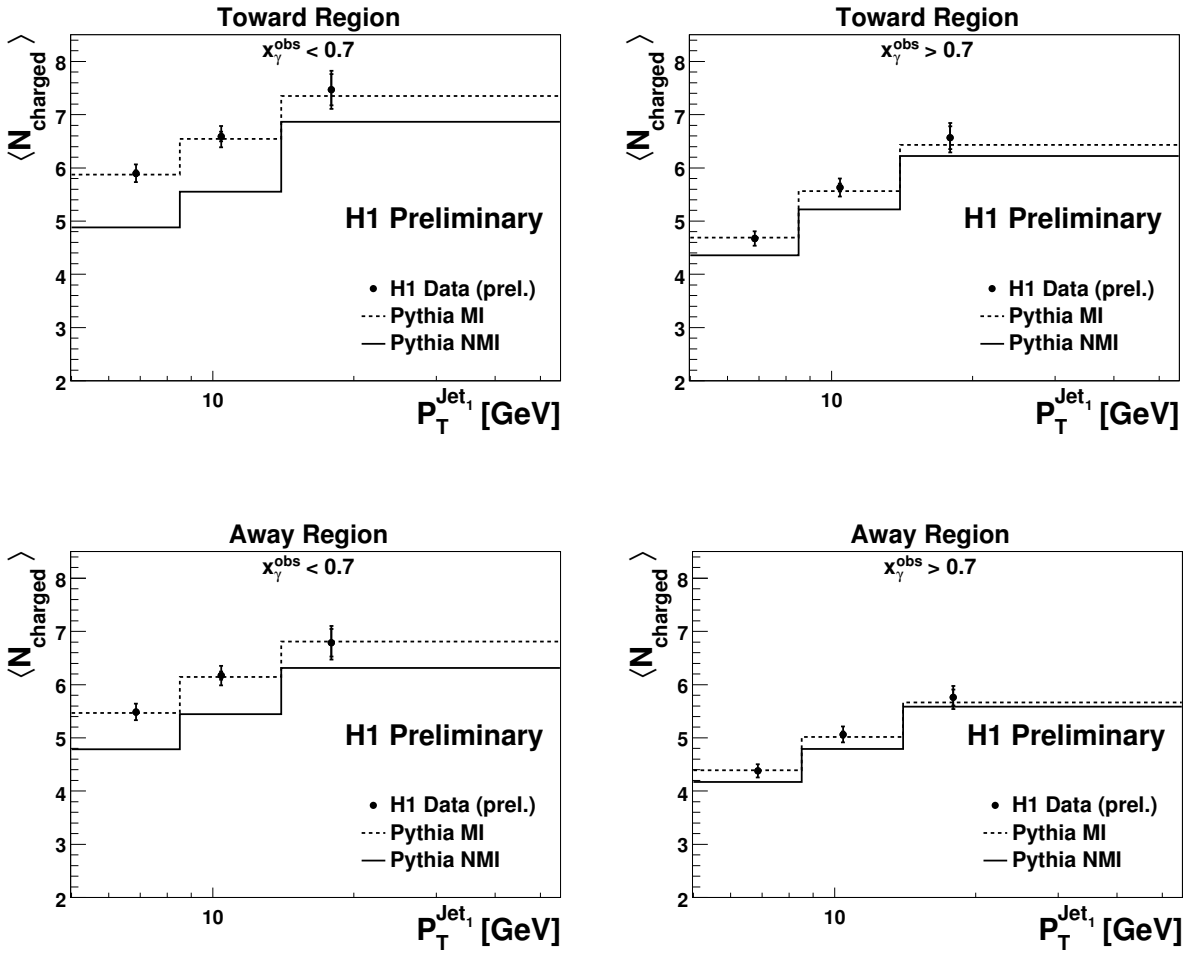


**Figure 6.2:** Average charged particle multiplicity as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$, for the toward and away regions (upper and lower plots, respectively) and for the low and high $x_\gamma^{obs}$ regions (left and right, respectively). Data are compared to PYTHIA with and without the MPI model.

low and high $x_\gamma^{obs}$ regions (left and right plots, respectively) and the high and low activity regions (upper and lower plots, respectively). The average charged particle multiplicity ranges between 0.5 and 2.5 particles depending on $x_\gamma^{obs}$, $P_T^{Jet_1}$ and the activity region.

In the high $x_\gamma^{obs}$ region, the differences between PYTHIA including MPI or PYTHIA with a single hard interaction, PYTHIA MI and PYTHIA NMI, respectively, are small. However, at low $P_T^{Jet_1}$ values many resolved photon events are present and therefore the contribution from MPI is still large as seen by the MC. This is especially visible in the high activity region.

For the low $x_\gamma^{obs}$ region, a larger average charged particle multiplicity than predicted from the simulation without MPI is needed. The simulation including MPI gives a reasonable description of the measurement over the full phase space region.

### 6.2.3 String Length Configurations

The possible colour connections between the partons become complex in presence of MPI (section 2.3). The effect of the different string length configurations is studied in this section.

In figure 6.4, the average track multiplicity as a function of $\Delta\phi^{Jet_1,\ h^\pm}$ is compared to the PYTHIA MPI model in two scenarios: short and long strings (the parameters and their description can be found in appendix D). In the $x_\gamma^{obs} > 0.7$ region, the differences between long and short strings are almost negligible, whereas in the low $x_\gamma^{obs}$ regions are larger (in the transverse regions up to 0.5 charged particles). While in the toward and away regions the differences are usually not very large between short and long strings, in the transverse regions the long string configuration is preferred.

Figure 6.5 shows the average charged particle multiplicity as a function of $P_T^{Jet_1}$ for the toward and away regions. The data are compared to the PYTHIA MPI model in two scenarios: short and long strings for the low and high $x_\gamma^{obs}$ regions. Again, the differences are small in the $x_\gamma^{obs} > 0.7$ region while they are larger for low $x_\gamma^{obs}$ values. Given the uncertainties in the toward and away regions, no scenario is preferred.

In the transverse regions, figure 6.6, the differences between short and long strings are large in the $x_\gamma^{obs} < 0.7$ region. Here, the long string configuration is preferred.

### 6.2.4 Higher Order Contributions

It is interesting to study whether the charged particle multiplicities could also be described by including more higher order QCD radiation. This can be approximately simulated by changing the upper scale in the initial and final parton shower (PS) simulation (section 1.6).

This was done in figure 6.7 by comparing the simulation provided by PYTHIA NMI vary-

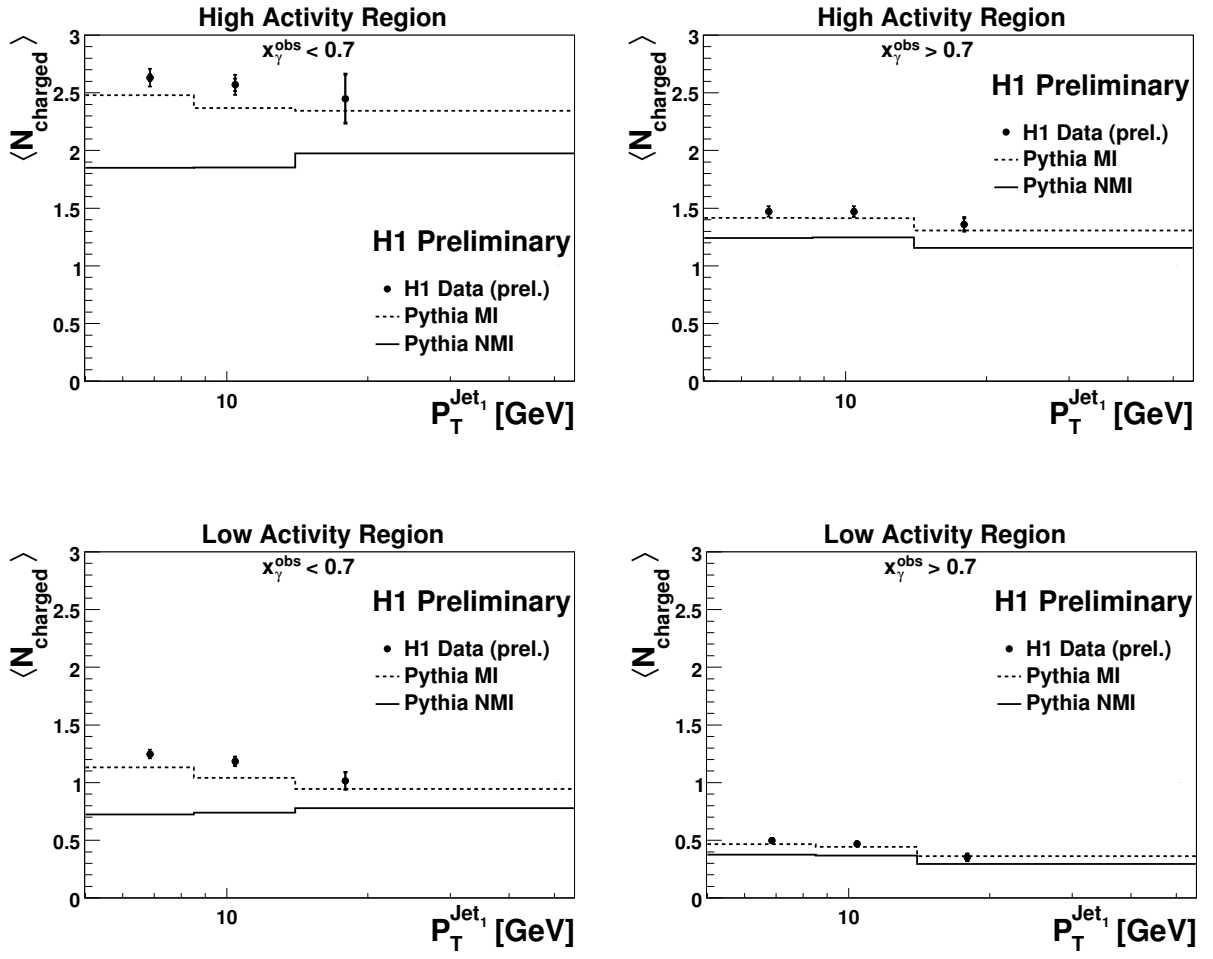


**Figure 6.3:** Average charged multiplicity in the transverse high and low activity regions (upper and lower plots, respectively) for the low and high $x_\gamma^{obs}$ regions (left and right, respectively). Data are compared to PYTHIA with and without MPI.

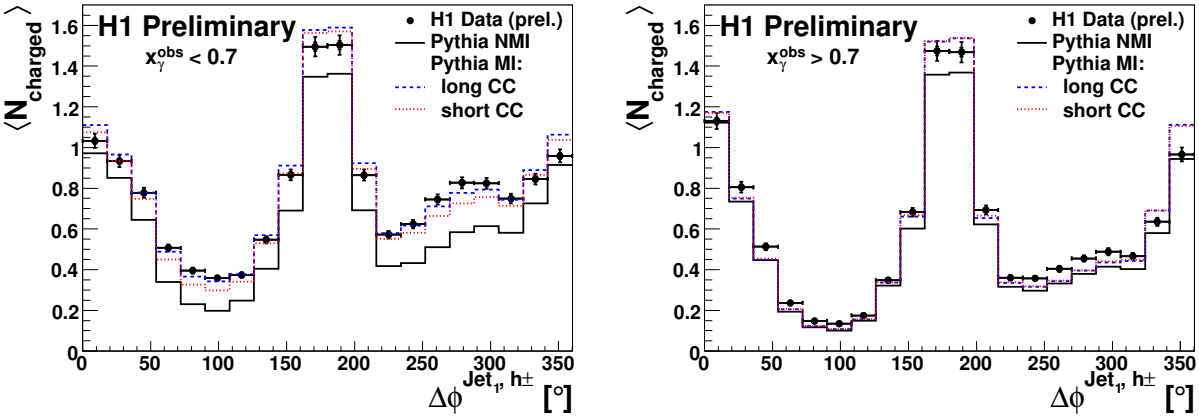


**Figure 6.4:** Average track multiplicity as a function of the azimuthal angle with respect the leading jet, $\Delta\phi^{Jet_1,\ h^\pm}$. Two $x_\gamma^{obs}$ regions are shown: a resolved photon enriched region with $x_\gamma^{obs} < 0.7$ (left) and a photon point-like enriched region with $x_\gamma^{obs} > 0.7$ (right). The effect of considering short and long strings in MPI is shown.

ing the upper scale for the PS between one and four times the hard scattering scale, fine dashed and dashed, respectively. The variation in the average charged particle multiplicity is shown for the transverse regions in the two $x_\gamma^{obs}$ regions. Although the difference between the upper scales allow for more parton radiation at higher scales, for high $x_\gamma^{obs}$ values the differences in the average charged particle are small. At low $x_\gamma^{obs}$ values the differences are larger, especially in the high activity region. However, the increase of the upper scale contributes at most 0.3 charged particles, which is clearly not enough to describe the data.

The higher order contributions can also be studied by simulating them using the CCFM parton evolution equations instead of DGLAP (section 1.6). This comparison is done in section 6.4.

## 6.3 Minijet Multiplicity as a function of $P_T^{Jet_1}$

In this section, the measurement of low transverse momentum jets (minijets) with $P_T^{Minijets} >$ 3.5 GeV and $-1.5 < \eta^{Minijets} < 2.79$ in dijet events with $P_T^{Jets} > 5$ GeV and $-1.5 < \eta^{Jets} < 2.79$ in the photoproduction regime $Q^2 < 0.01$ GeV$^2$ is presented.

In figure 6.8, the average minijet multiplicity is shown as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$, for the toward and away regions (upper and lower plots, respectively) and for the low and high $x_\gamma^{obs}$ regions (left and right, respectively). Since in the toward region the leading jet is always present, the average number of minijets

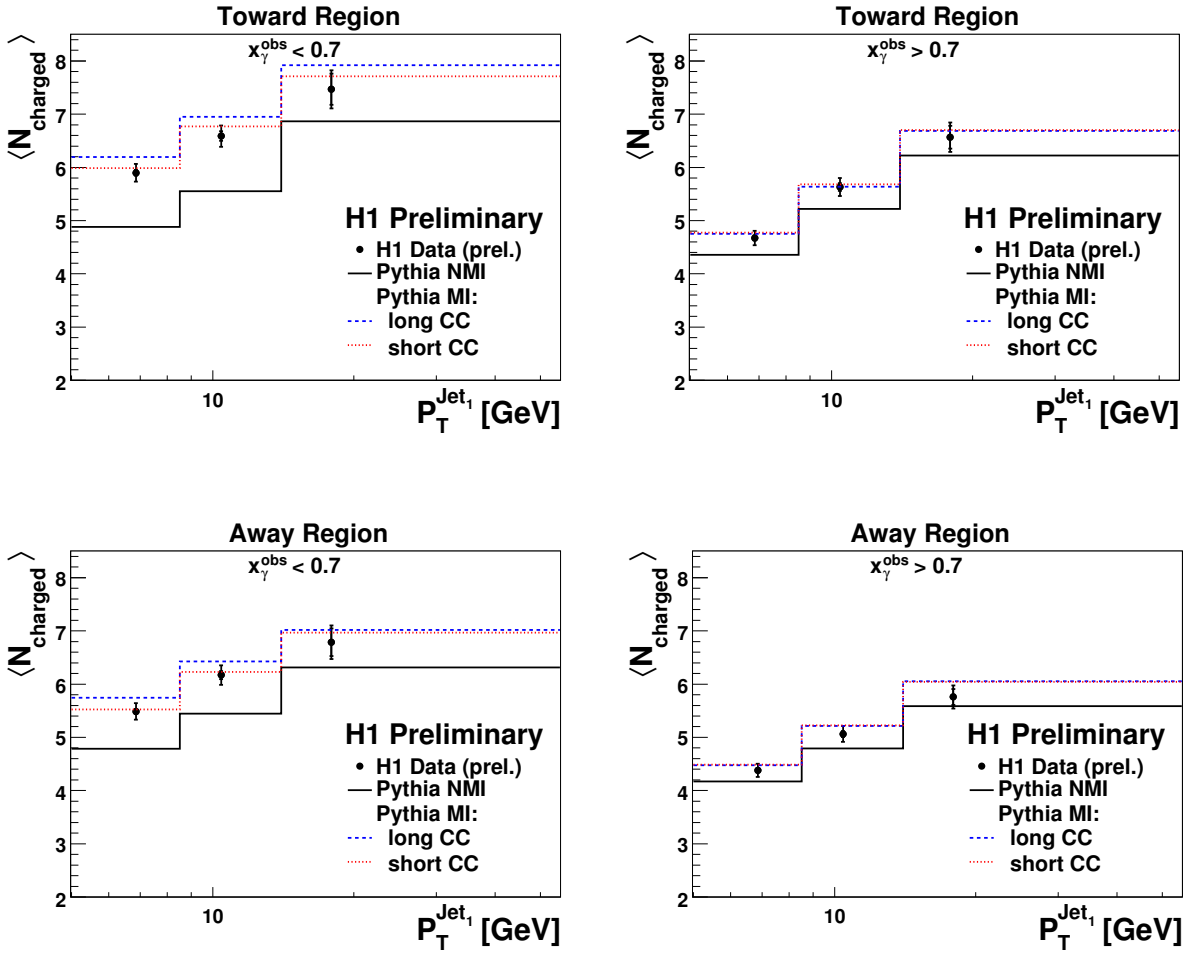


**Figure 6.5:** Average charged particle multiplicity as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$, for the toward and away regions (upper and lower plots, respectively) and for the low and high $x_\gamma^{obs}$ regions (left and right, respectively). The effect of considering short and long strings in MPI is shown.

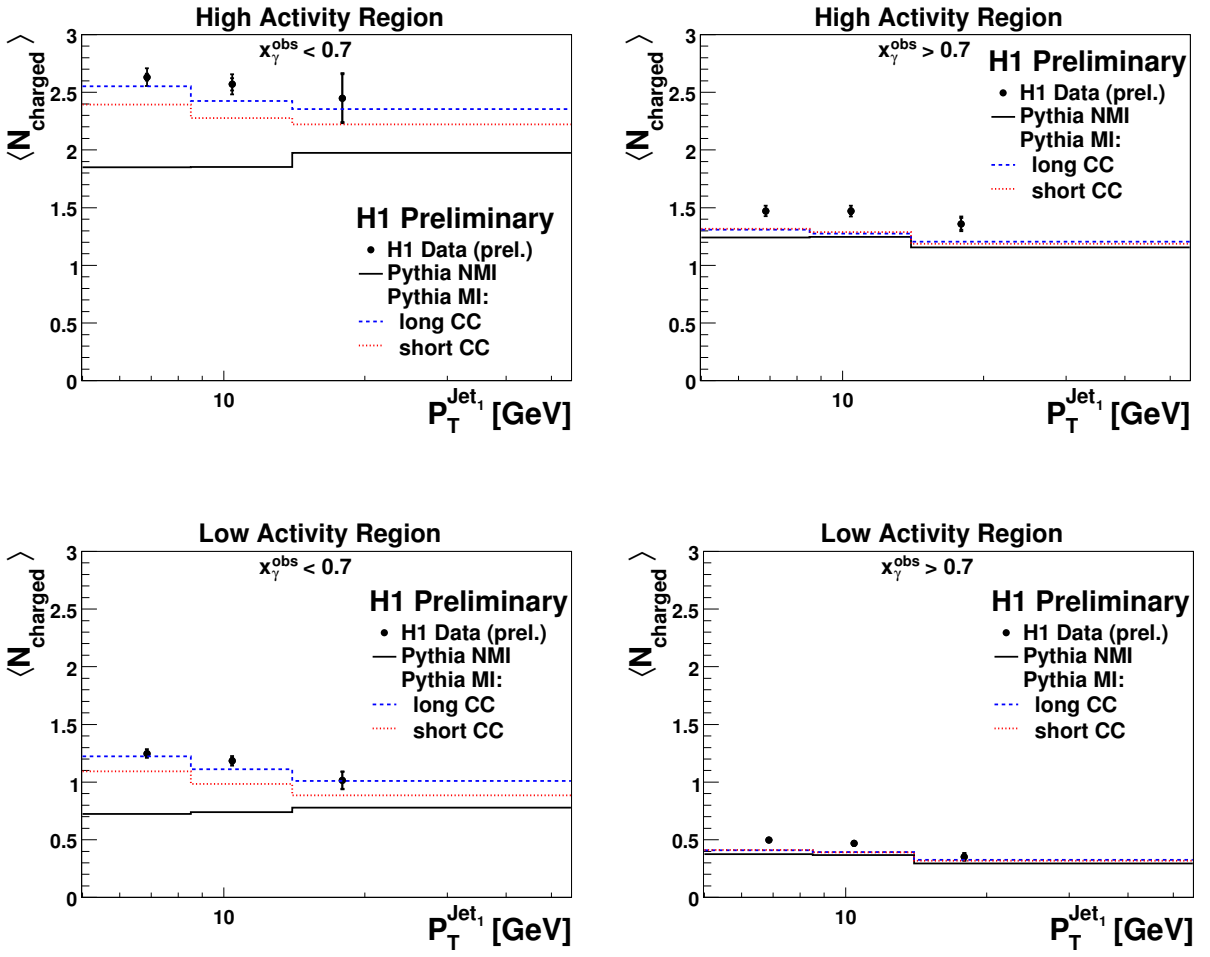


**Figure 6.6:** Average charged multiplicity in the transverse high and low activity regions (upper and lower plots, respectively) for the low and high $x_\gamma^{obs}$ regions (left and right, respectively). The effect of considering short and long strings in MPI is shown.

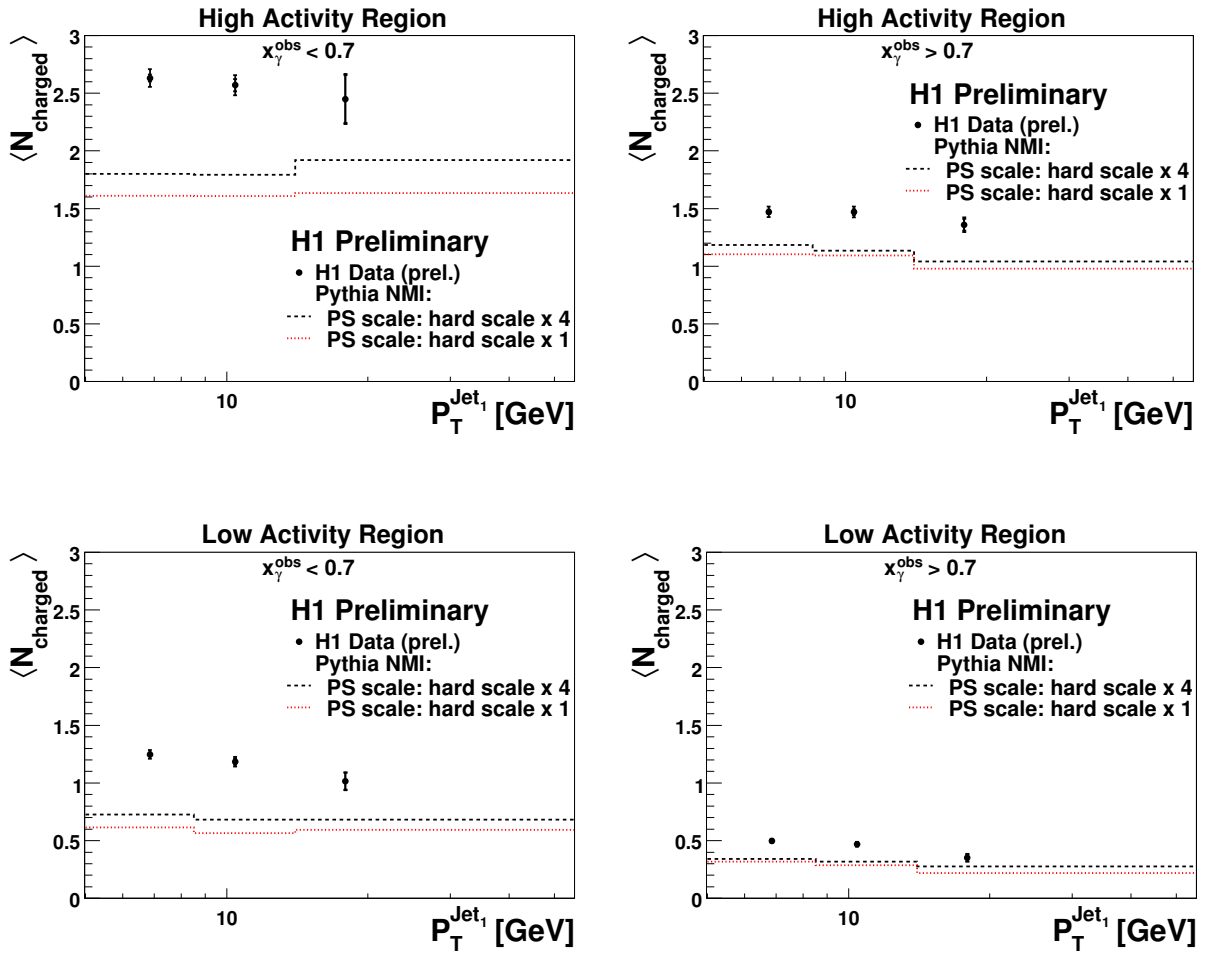


**Figure 6.7:** The variation in the average charged particle multiplicity for the transverse regions in the two $x_\gamma^{obs}$ regions is shown when varying the upper scale for the parton showers between one and four times the hard scattering scale.

is slightly above unity. In the low $x_\gamma^{obs}$ region the average is roughly constant around 1.1 minijets, while in the high $x_\gamma^{obs}$ region the average is closer to 1. In the away region the average number of minijets grows with $P_T^{Jet_1}$ from 1 to 1.4.

In the region $x_\gamma^{obs} > 0.7$ the simulation containing only one hard interaction (PYTHIA NMI) and with MPI (PYTHIA MI) give similar predictions and describe the data well. In the low $x_\gamma^{obs}$ region the simulations differ and the simulation including MPI describes the data better.

The average minijet multiplicity in the high and low activity transverse regions (upper and lower plots, respectively) and for the $x_\gamma^{obs} < 0.7$ and the $x_\gamma^{obs} > 0.7$ regions (left and right plots, respectively) are presented in figure 6.9.

In the high $x_\gamma^{obs}$ region, PYTHIA MI and PYTHIA NMI give in general similar predictions. However, the resolved events still present in this region allow MPI which enhances the average minijet multiplicity in this region and improves the description. In the low $x_\gamma^{obs}$ region, the average minijet multiplicity is larger for the simulation including MPI. In contrast to the case of the charged particle multiplicity, the simulation including MPI cannot describe the data.

In general, the differences between both simulations decrease with increasing $P_T^{Jet_1}$ but the $P_T^{Jet_1}$ dependence is different. In the high activity region, while the average number of minijets grows for PYTHIA NMI from 0.12 to 0.18, for PYTHIA MI they are roughly constant as a function of $P_T^{Jet_1}$ around 0.23. The data are roughly constant around 0.37 with a variation of around 0.2 minijets in average.

The measured average number of minijets suggests that the number of MPI could be larger. To increase the average number of MPI per event the $p_{\perp,min}$ is decreased[1](section 2.3.1). If the average number of MPI increases the average number of charged particles increases too. Since the average number of charged particles is well described, a mechanism to keep the particle production while increasing the number of jets is needed. This can be achieved if the colour strings between MPI is minimized, i.e. less particle production, while the average number of MPI is increased, i.e. more minijets are produced.

In figure 6.10 a simulation with a lower $p_{\perp,min}$ value was used with short strings. The average minijet multiplicity in the toward and away regions are shown. Even with an enhanced number of MPI the average minijet multiplicity is well described in these regions.

In the transverse regions, the effect of increasing the average number of MPI is more visible in the average minijet multiplicity. In figure 6.11 it is shown how for this simulation the average minijet multiplicity increases in all the transverse regions and the description of the data is improved although not perfect.

This shows that the average charged particle and the minijet multiplicities are sensitive to how the fragmentation process is done when MPI occur and the average number of MPI. Therefore, these observables can be used to understand MPI. Within the PYTHIA

[1]It was changed from the default value $p_{\perp,min} = 1.9$ GeV to $p_{\perp,min} = 1.5$ GeV (appendix D).

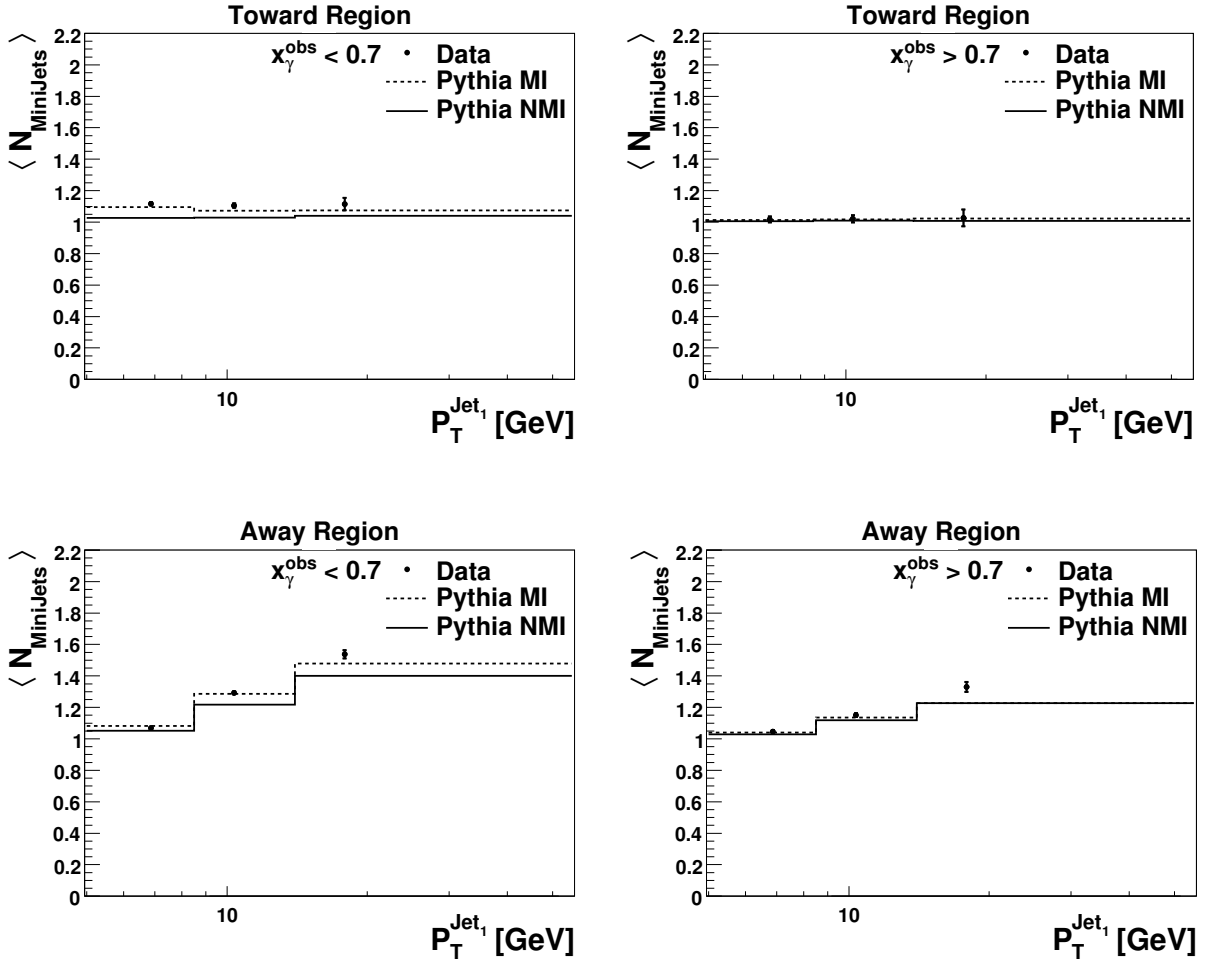


**Figure 6.8:** Average minijet multiplicity as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$ for the toward and away regions (upper and lower plots, respectively) and for the low and high $x_\gamma^{obs}$ regions (left and right, respectively). Data are compared to PYTHIA with and without MPI.

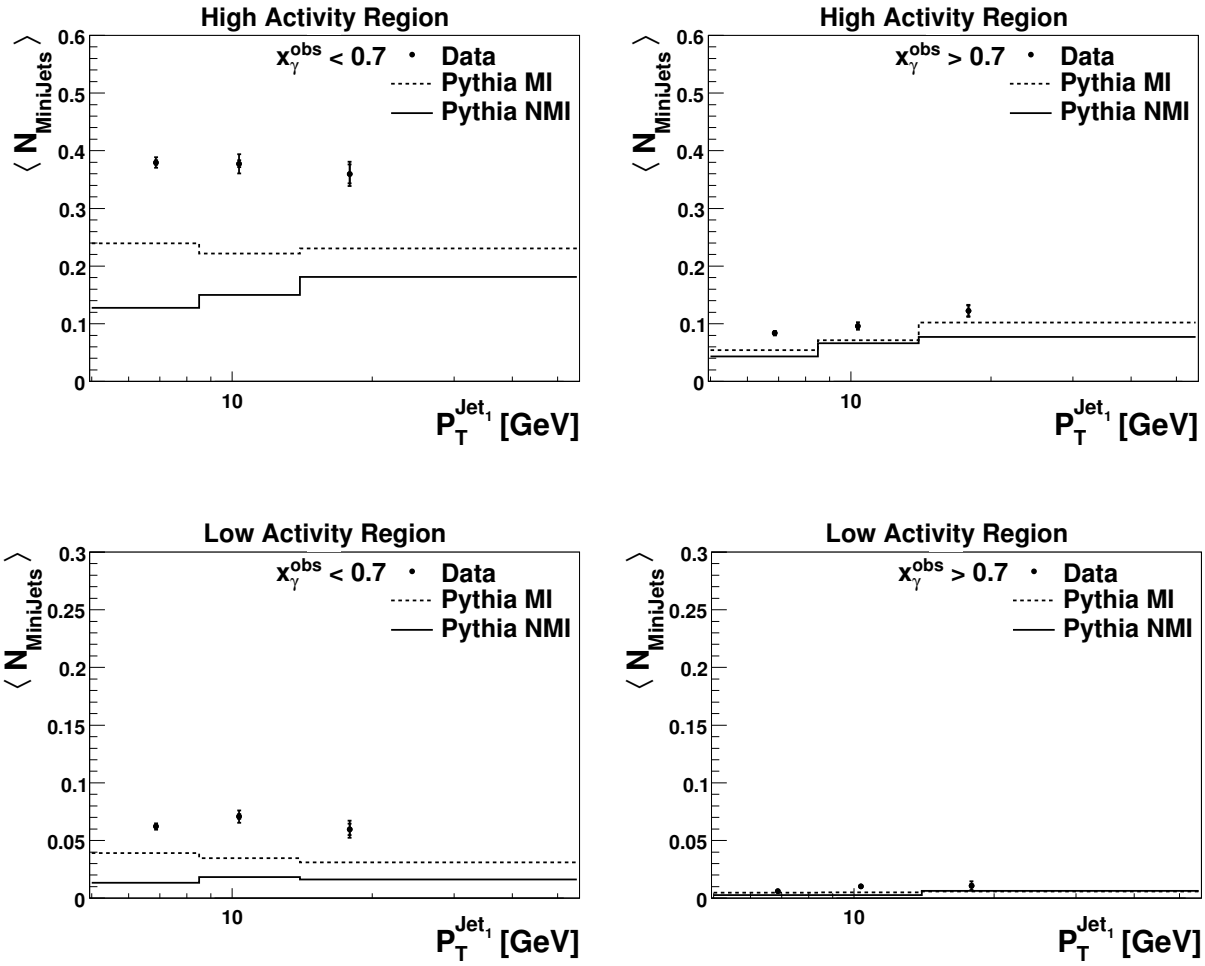


**Figure 6.9:** Average minijet multiplicity as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$, in the high and low activity transverse regions (upper and lower plots, respectively) and for the $x_\gamma^{obs} < 0.7$ and the $x_\gamma^{obs} > 0.7$ regions (left and right plots, respectively). Data are compared to PYTHIA with and without MPI.

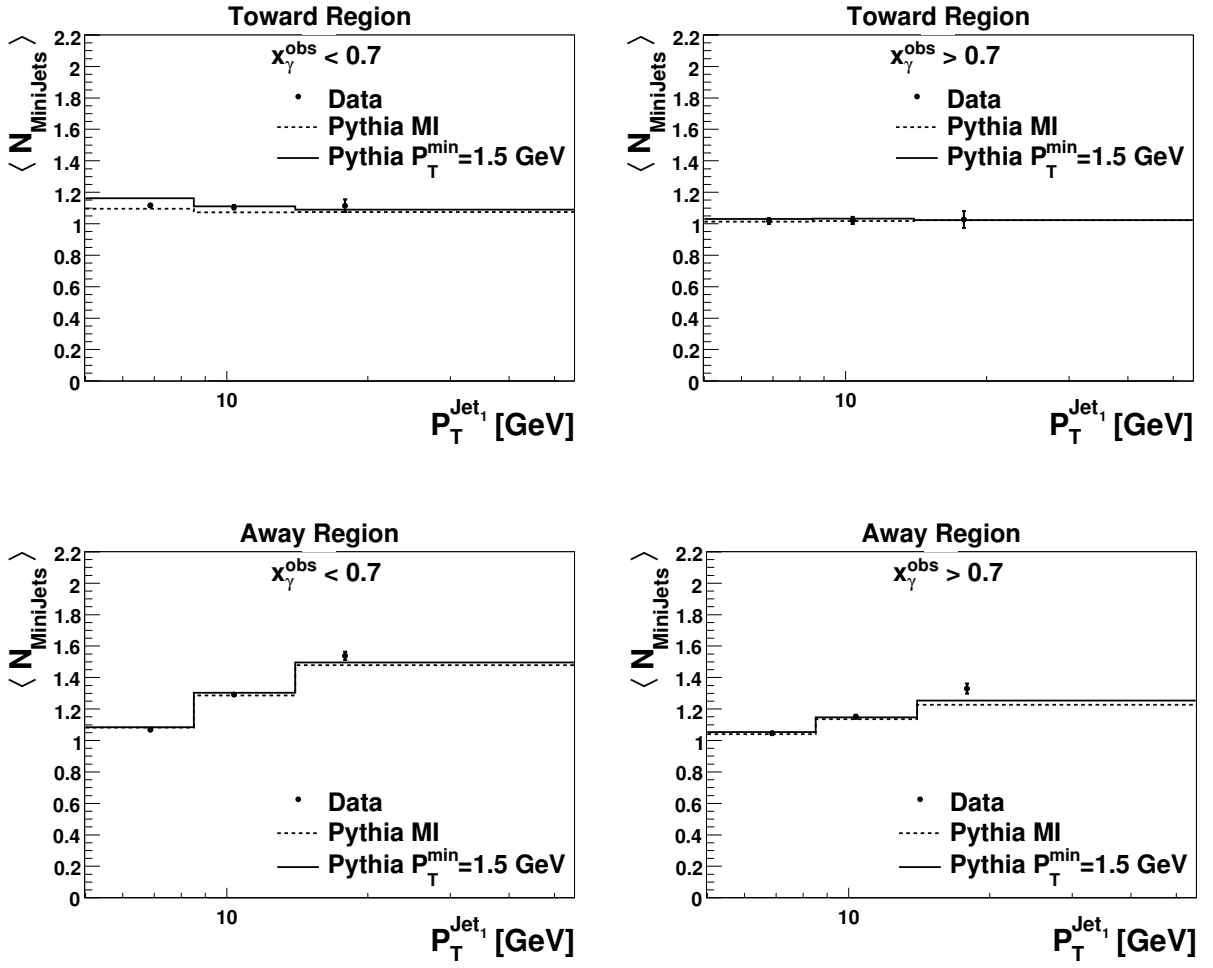


**Figure 6.10:** Average charged particle multiplicity as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$, for the toward and away regions (upper and lower plots, respectively) and for the low and high $x_\gamma^{obs}$ regions (left and right distributions, respectively). Data are compared to PYTHIA with an enhanced average number of MPI and short colour strings.

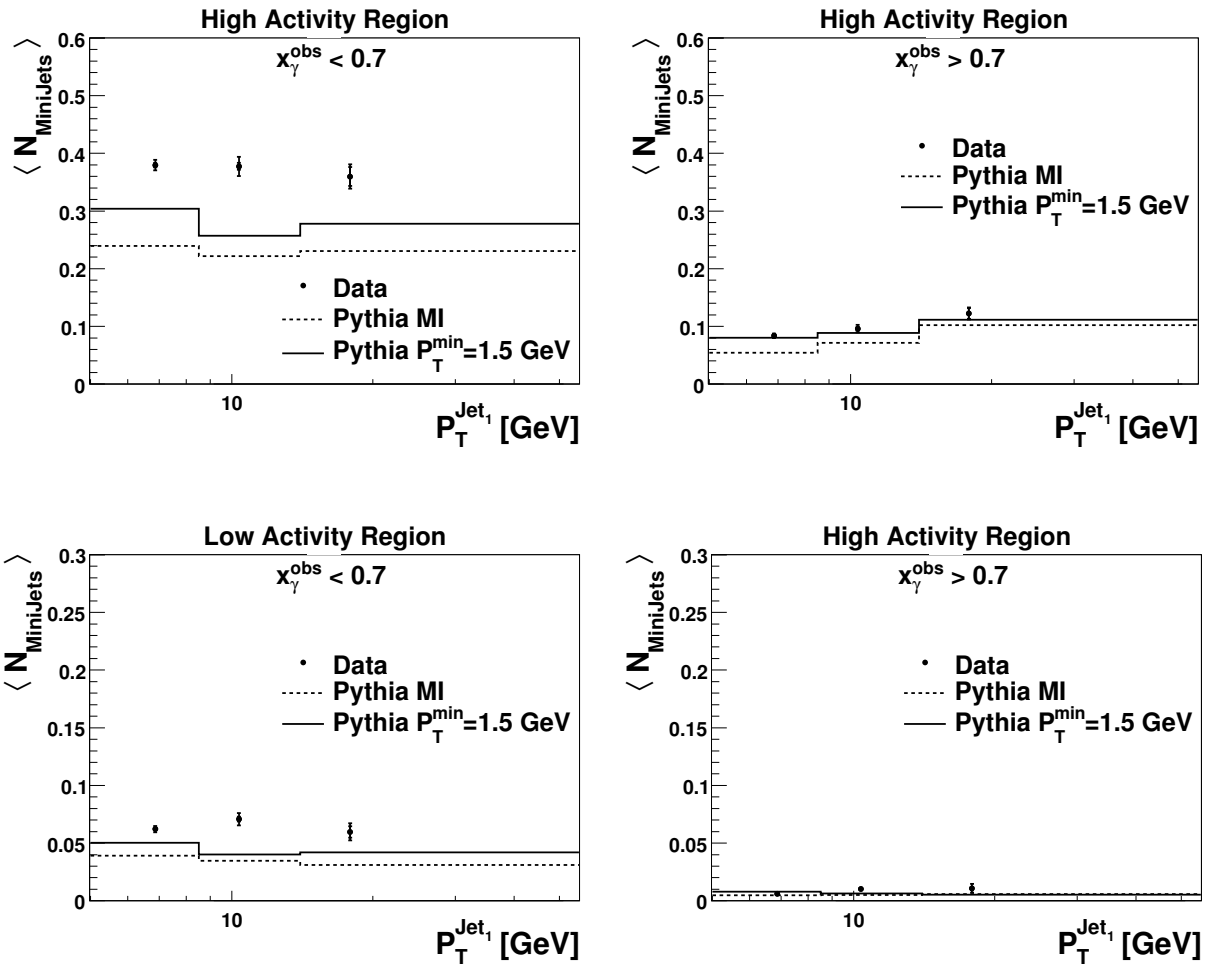


**Figure 6.11:** Average minijet multiplicity as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$, in the high and low activity transverse regions (upper and lower plots, respectively) and for the $x_\gamma^{obs} < 0.7$ and the $x_\gamma^{obs} > 0.7$ regions (left and right plots, respectively). Data are compared to PYTHIA with an enhanced average number of MPI and short colour strings.

model this means that these observables are sensitive to the length of the colour strings and the number of MPI. In future steps, the PYTHIA model parameters can be fitted to these observables for a better description of the data using a program like [70]

## 6.4 Model Comparisons

It is important to compare the data to different parton evolution equations and fragmentation models in order to have a proper interpretation of the measured average charged particle and minijet multiplicities. The data are compared in this section to two simulations: CASCADE[2] with CCFM parton evolution equations and the Lund string model and HERWIG/JIMMY with DGLAP parton evolution equations, the cluster fragmentation model and a model for MPI.

In figure 6.12, the average charged particle multiplicity in the toward and away regions for the low and high $x_\gamma^{obs}$ regions is shown. While CASCADE describes all regions HERWIG generally underestimates the data.

The average charged particle multiplicity in the transverse high and low activity regions and for low and high $x_\gamma^{obs}$ values is shown in figure 6.13. At high $x_\gamma^{obs}$ CASCADE describes data both in the high and the low activity regions. In these regions HERWIG underestimates particle multiplicity.

For the high activity region at low $x_\gamma^{obs}$, the description given by CASCADE differs from data at low transverse momentum of the leading jet, $P_T^{Jet_1}$. At high $P_T^{Jet_1}$ the CASCADE description is closer to data. In the low activity region and low $x_\gamma^{obs}$ values, the charged particle multiplicity predicted by CASCADE is too low. The rather good description given by CASCADE in the high activity region and the underestimation given by CASCADE in the low activity region can be explained because the contribution from higher orders is expected to be larger in the high activity region than in the low activity region (section 2.4). Since CASCADE is expected to include part of the NLO corrections (section 1.6) it is expected that CASCADE describes data in the high activity region although not perfectly since CASCADE does not include contributions from resolved photons.

HERWIG does not describe data neither in the high nor in the low activity regions, for low or high $x_\gamma^{obs}$ values.

In figure 6.14 the average minijet multiplicity in the toward and away regions for the low and high $x_\gamma^{obs}$ regions is shown. CASCADE and HERWIG describe the toward region for both high and low $x_\gamma^{obs}$ values. In the away region CASCADE describes the data but HERWIG cannot describe the increase of the minijet multiplicity as a function of $P_T^{Jet_1}$.

The average minijet multiplicity in the transverse regions is shown in figure 6.15 where the data are compared to CASCADE and HERWIG. In the high $x_\gamma^{obs}$ region, neither CASCADE nor HERWIG describe the high and low transverse regions. In the low $x_\gamma^{obs}$

[2]The uPDF used in this simulation is the CCFM set A0.

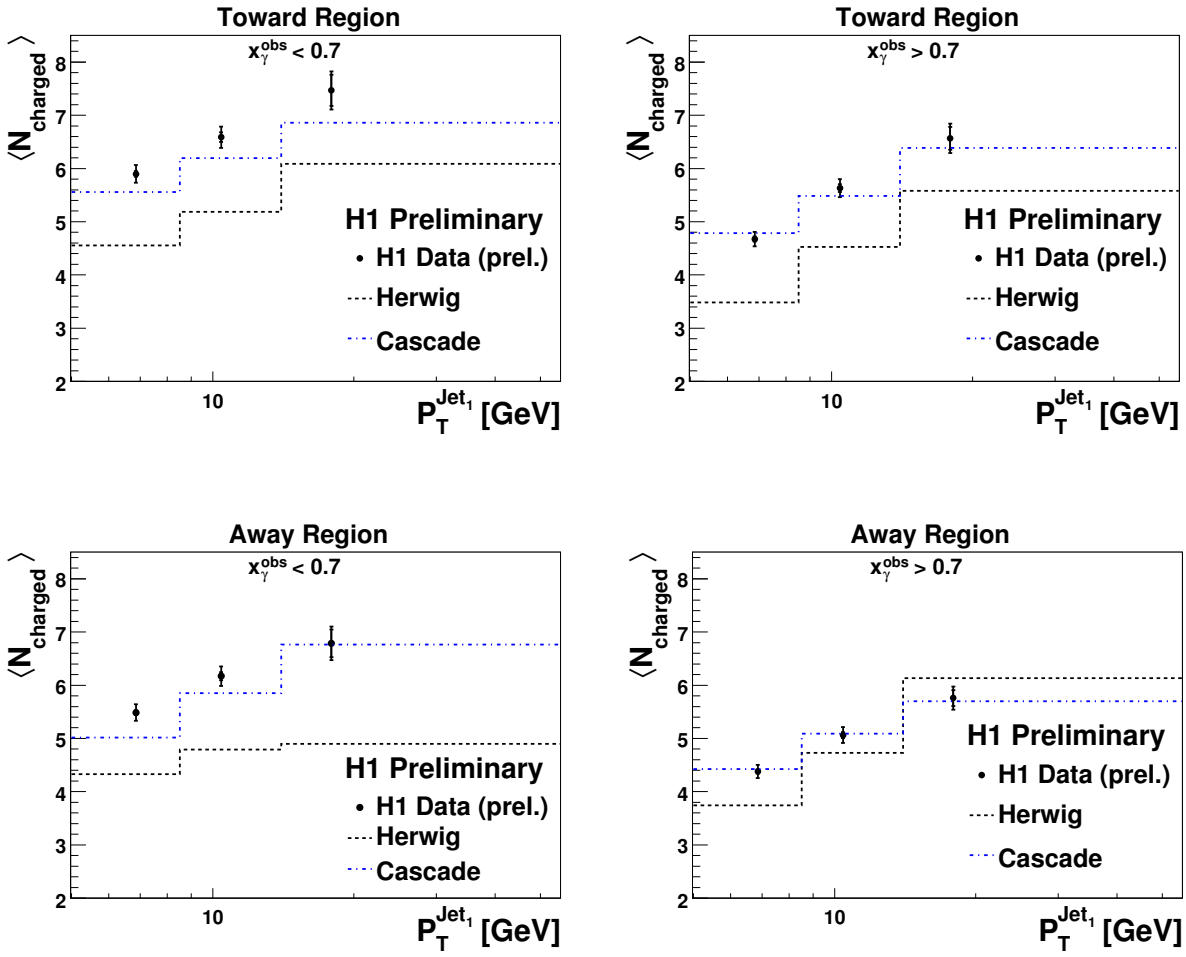


**Figure 6.12:** Average charged particle multiplicity as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$, for the toward and away regions (upper and lower plots, respectively) and for the low and high $x_\gamma^{obs}$ regions (left and right distributions, respectively). Data are compared to CASCADE and HERWIG/JIMMY.

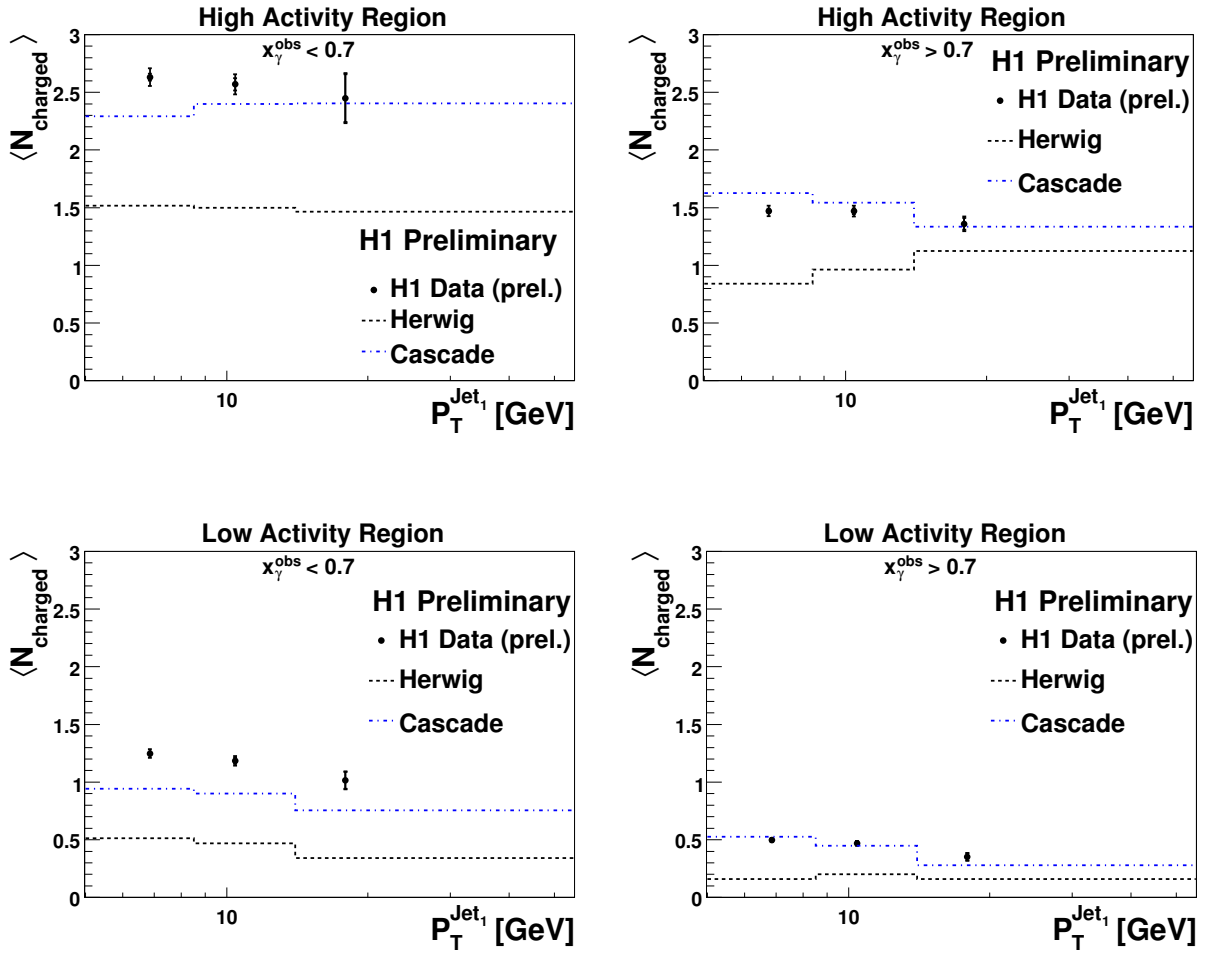


**Figure 6.13:** Average charged particle multiplicity in the transverse high and low activity regions (upper and lower plots, respectively) for the low and high $x_\gamma^{obs}$ regions (left and right, respectively). Data are compared to CASCADE and HERWIG/JIMMY.

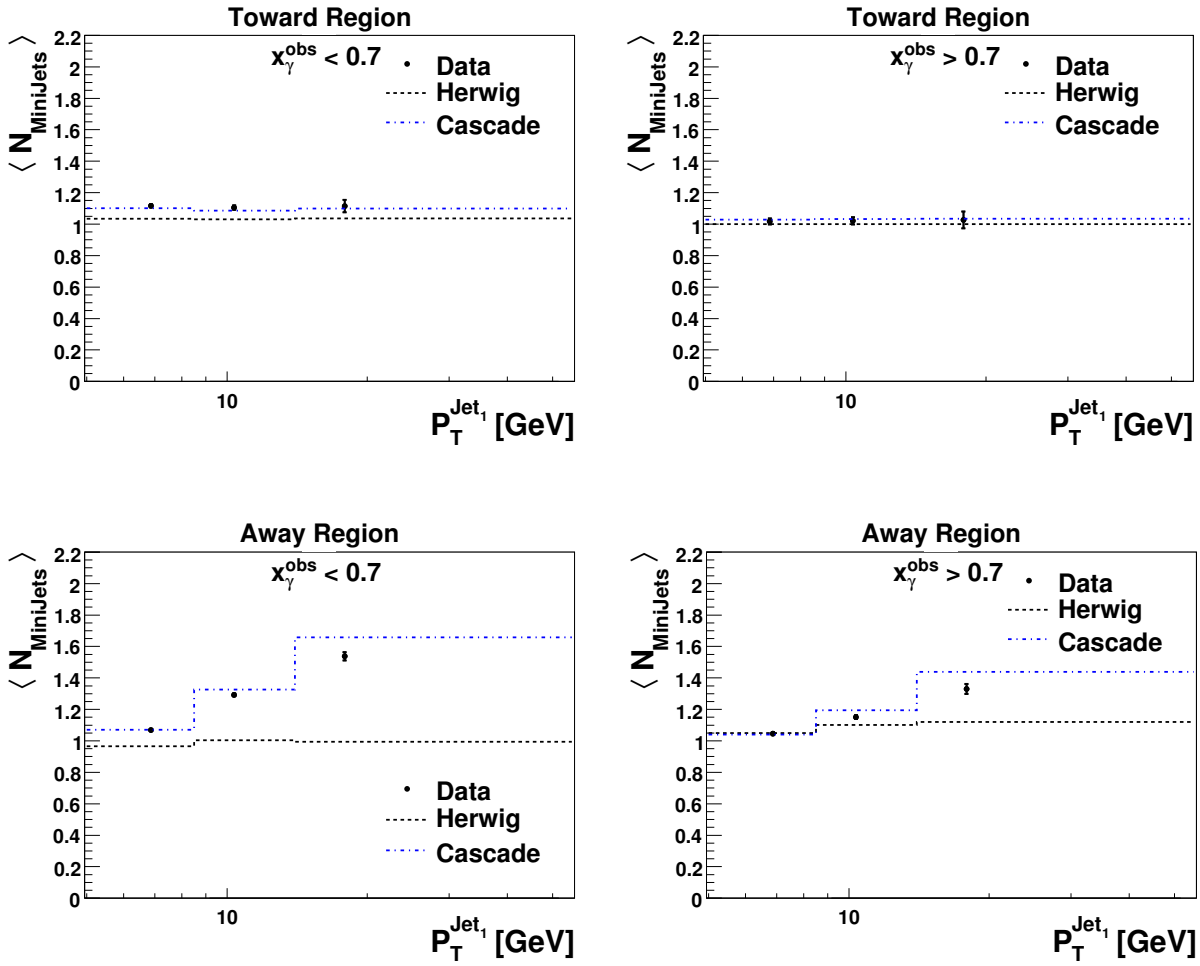


**Figure 6.14:** Average minijet multiplicity as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$, for the toward and away regions (upper and lower plots, respectively) and for the low and high $x_\gamma^{obs}$ regions (left and right, respectively). Data are compared to CASCADE and HERWIG/JIMMY.

region, CASCADE is close to the data while the number of minijets in HERWIG is too low.

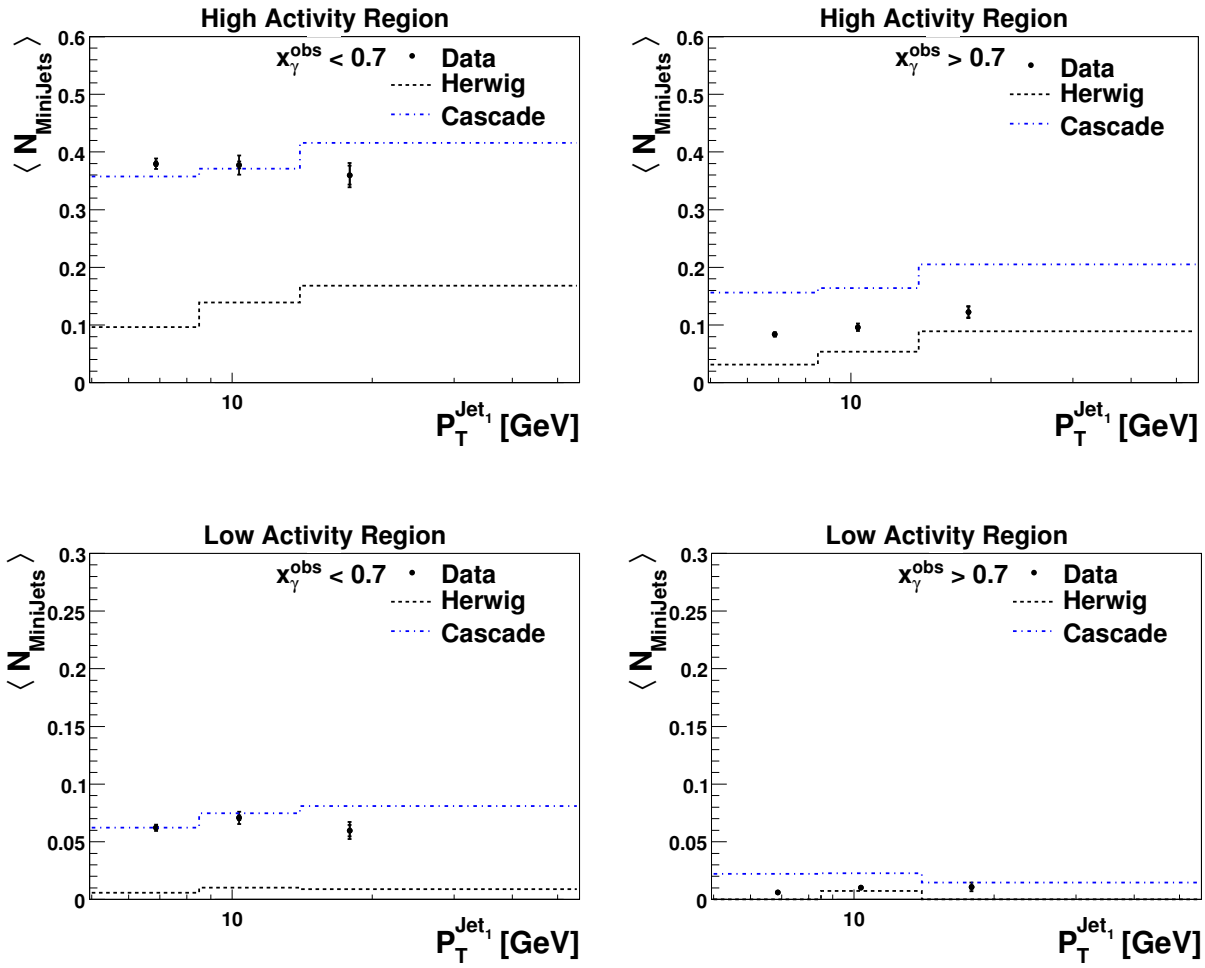


**Figure 6.15:** Average minijet multiplicity in the transverse high and low activity regions (upper and lower plots, respectively) for the low and high $x_\gamma^{obs}$ regions (left and right, respectively). Data are compared to CASCADE and HERWIG/JIMMY.

# Chapter 7

# Conclusions and Outlook

The average charged particle and low transverse momentum jet, minijet, multiplicity in dijet photoproduction events has been measured as a function of the azimuthal angle with respect to the leading jet. The measurement was compared to PYTHIA, CASCADE and HERWIG. The charged particle and minijet multiplicity has been investigated for two different regions in $x_\gamma^{obs}$, separating the point-like from the hadronic (resolved) photon and as a function of the leading jet transverse momentum.

In the region of $x_\gamma^{obs} > 0.7$, the average **charged particle multiplicity** is reasonably well described by PYTHIA including one hard scattering process supplemented with parton showers and hadronization and without MPI (PYTHIA NMI) and CASCADE. PYTHIA with the multiple parton interactions (MPI) model (PYTHIA MI) also describes the charged particle multiplicity.

The region of $x_\gamma^{obs} < 0.7$ shows that the charged particle multiplicity is larger than what is expected from PYTHIA NMI, and is well described when MPI are included, i.e PYTHIA MI. CASCADE describes the average charged particle multiplicity in the transverse high activity region for high $P_T^{Jet_1}$ values but not the low activity region.

It was also shown that the charged particle multiplicity in the transverse region cannot be accounted by increasing the amount of initial and final state radiation. The CCFM parton evolution equation does not describe the data in all regions. However, the prediction given by CASCADE points to the fact that a better account for the parton kinematics can improve the particle production description.

In the high $x_\gamma^{obs}$ region, the **average minijet multiplicity** is well described by PYTHIA with and without MPI in the toward and away regions. The transverse regions are better described with PYTHIA MI. CASCADE describes perfectly the toward and away regions but overestimates the minijet production in the transverse regions.

The toward and away regions are well described by PYTHIA NMI and CASCADE in the low $x_\gamma^{obs}$ region. PYTHIA MI also described the data well in these regions. In the transverse regions, PYTHIA NMI underestimates the data. The prediction given by

PYTHIA MI is better but in the current simulation this is not enough to describe data. In contrast, the description given by CASCADE is very close to data.

Although no simulation describes all regions properly, it was shown that: a) increasing the average number of MPI per event increases the average charged particle multiplicity, b) the long string lengths were preferred for the average charged particle measurement and c) increasing the average number of MPI per events improves the minijet multiplicity. This shows that the data could probably be described by using short string lengths with a larger average of MPI per event.

To understand MPI it is necessary to understand better the $x_\gamma^{obs} > 0.7$ region. In this region, it was shown by CASCADE that we still need to improve the description given by the parton showers. In this $x_\gamma^{obs}$ region no MPI are present and we should be able to describe the data with pure perturbative calculations. Here, *ep* data offers a unique clean environment. This allows to understand the initial and final parton showers, constituents of the underlying event[1](UE). Once this is understood, and using the same observables, the low $x_\gamma^{obs}$ region can be used to study MPI. For this purpose, the data presented in the present analysis and in [60], where a similar strategy was used in DIS are very useful.

It is interesting to observe that the data in this analysis points to a possible preference for short string configurations, similar to measurements at TeVatron [71].

Since the HERWIG simulation gives a poor description of data, it is not possible to draw any conclusion about the different fragmentation and MPI models.

In conclusion, the measured multiplicities provide a set of observables which represents a valuable information to study the mechanism of MPI and the UE in general. This is possible since the observables are measured for the direct and resolved photon components which allows to separate the parton showers from MPI. This separation allows a very detailed investigation of the UE at HERA.

After validation of the parton showers description in the high $x_\gamma^{obs}$ region we will be in the position of estimating the average MPI per event. Nevertheless, we expect that once the description of parton showers is understood, the amount of average MPI per event is less than the predicted by the current PYTHIA model.

Appart from the question of understanding the fundamental interactions, the understanding of MPI is very important for the Large Hadron Collider (LHC) and cannot be done alone at hadron accelerators like at TeVatron. The Higgs[2]and the beyond the standard model searches will be affected by the underlying event and therefore it is important to understand its mechanism.

[1]The underlying event (UE) is defined as everything except the lowest order process (section 2.2)
[2]For example in WH or ZH production [72].

# Appendix A

# Derivation of $\mathbf{x}_\gamma^{\mathrm{obs}}$

In this appendix it will be shown how to calculate $x_\gamma^{\mathrm{obs}}$ from hadrons or detector level quantities and neglecting multiple parton interactions.

If $x_\gamma$ and $x$ are the photon and proton energy fraction entering into the hard interaction then by energy conservation:

$$x_\gamma E_\gamma + x E_\mathrm{p} = \sum_i^{\mathrm{partons}} E_i, \tag{A.1}$$

where $E_\gamma$ and $E_\mathrm{p}$ are the photon and proton energies, respectively, the sum runs over all partons from the hard scattering (see figure A.1) and $E_i$ the energy of the $i$th parton. Furthermore, by momentum conservation:

$$x|p_\mathrm{p}|\cos\theta_\mathrm{p} - x_\gamma|p_\gamma|\cos\theta_\gamma = \sum_i^{\mathrm{partons}} |p_i|\cos\theta_i = \sum_i^{\mathrm{partons}} P_{z,i} \tag{A.2}$$

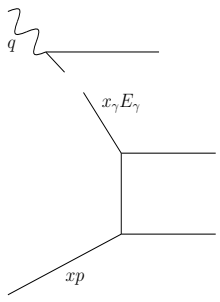


**Figure A.1:** ep scattering

considering the photon and the proton to be collinear and $\theta_{\gamma,\mathrm{p}} = 0$:

$$xE_{\mathrm{p}} - x_\gamma E_\gamma = \sum_i^{\text{partons}} P_{z,i}. \tag{A.3}$$

Now, subtracting equation A.1 from A.2:

$$2x_\gamma E_\gamma = \sum_i^{\text{partons}} (E_i - P_{z,i})\,. \tag{A.4}$$

From the definition of $y$, the inelasticity, in section 1.1 and using the Jacquet-Blondel method $y \simeq y_{JB} = \frac{\sum_{f\neq e}(E_f - p_{z,f})}{2E_e}$, where the sum runs over all particles[1] in the final state except the electron:

$$y = \frac{E_\gamma}{E_{\mathrm{e}}} = = \frac{\sum_{f\neq e}(E_f - p_{z,f})}{2E_e} \Longrightarrow 2E_\gamma = \sum_{f\neq e}(E_f - P_{z,f}). \tag{A.5}$$

Combining A.4 and A.5

$$2x_\gamma E_\gamma = \sum_i^{\text{partons}} (E_i - P_{z,i}) = x_\gamma \sum_f E_i - p_{z,f}$$

and solving for $x_\gamma$:

$$x_\gamma = \frac{\sum_i^{\text{partons}} (E_i - P_{z,i})}{\sum_f E_f - p_{z,f}}. \tag{A.6}$$

We now show the derivation of $y_{JB}$. The four-momenta are defined as:

$$\begin{aligned} \text{proton } 4-\text{momentum} : p &= (E_{\mathrm{p}}, E_{\mathrm{p}}, \vec{0}) \\ \text{electron } 4-\text{momentum} : k &= (E_{\mathrm{e}}, -E_{\mathrm{e}}, \vec{0}) \\ \text{scattered electron } 4-\text{momentum} : k' &= (E_{e'}, E_{e'}\cos\theta_e, E_{e'}\sin\theta_e) \end{aligned}$$

[1] Notice that here the sum runs over all particles and not only from the hard scattering as will become clear later in derivation of $y_{JB}$.

Using the definition of $y$ we obtain:

$$y = \frac{p\ q}{p\ k} = 1 - \frac{E_{e'}}{2E_e}(1 - \cos\theta_e). \tag{A.7}$$

This can be re-written using the sum:

$$\sum_f (E_f - p_{z,f}) = \sum_f E_f(1 - \cos\theta_f) \simeq \sum_{i=e,p} E_i - p_{z,i} = 2E_e, \tag{A.8}$$

where the sum over $f$ runs over all the particles (electron included) and in the last step we used again the four-momenta definition of the electron and proton. From this equation and splitting the sum over $f$ into $\sum_f (E_f - p_{z,f}) = \sum_{f\neq e}(E_f - p_{z,f}) + (E_{e'} - p_{z,e'})$ it follows that

$$\frac{\sum_{f\neq e}(E_f - p_{z,f})}{2E_e} = 1 - \frac{E_{e'}}{2E_e}(1 - \cos\theta_e) = y. \tag{A.9}$$

This is the equation used for $y_{JB}$ in A.6. Finally, associating the partons from the hard scattering with the jets after hadronization:

$$x_\gamma^{\text{obs}} = \frac{\sum_i^{\text{N}_{\text{jets}}} (E_i - P_{z,i})}{\sum_f E_f - p_{z,f}}. \tag{A.10}$$

# Appendix B

# Measured Charged Particles Multiplicities

In this appendix the measured average charged particle multiplicity values are given with the uncertainties. Two uncertainties are given: from the unfolding method which includes statistical and from the uncertainties from the smearing matrix ($\Delta_{UF}$), and the systematic uncertainties ($\Delta_{Sys}$) which includes the hadronic energy scale, the tracking and the transverse activity resolution uncertainties. The uncertainties are added in quadrature. The total uncertainty ($\Delta_{tot}$) is also given in the tables.

In the charged particle multiplicity as a function of the jet with the highest transverse momentum, $P_T^{Jet_1}$, the average value a detector level is taken and not the bin centre.

| $P_T^{Jet_1}$ [GeV] | $\langle N_{ch} \rangle$ | $\Delta_{UF}$ | $\Delta_{Sys}$ | $\Delta_{tot}$ |
|---|---|---|---|---|
| 6.9 | 5.90 | 0.05 | 0.16 | 0.17 |
| 10.4 | 6.59 | 0.09 | 0.18 | 0.20 |
| 17.9 | 7.47 | 0.29 | 0.21 | 0.36 |

**Table B.1:** Charged particle average in the **toward** region for events with $x_\gamma^{obs} < 0.7$ values as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$.

| $P_T^{Jet_1}$ [GeV] | $\langle N_{ch} \rangle$ | $\Delta_{UF}$ | $\Delta_{Sys}$ | $\Delta_{tot}$ |
|---|---|---|---|---|
| 6.9 | 4.67 | 0.05 | 0.13 | 0.14 |
| 10.4 | 5.63 | 0.07 | 0.15 | 0.17 |
| 17.9 | 6.57 | 0.21 | 0.18 | 0.28 |

**Table B.2:** Charged particle average in the **toward** region for events with $x_\gamma^{obs} > 0.7$ values as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$.

| $P_T^{Jet_1}$ [GeV] | $\langle N_{ch} \rangle$ | $\Delta_{UF}$ | $\Delta_{Sys}$ | $\Delta_{tot}$ |
|---|---|---|---|---|
| 6.9 | 5.49 | 0.04 | 0.15 | 0.15 |
| 10.4 | 6.17 | 0.08 | 0.17 | 0.18 |
| 17.9 | 6.79 | 0.26 | 0.18 | 0.32 |

**Table B.3:** Charged particle average in the **away** region for events with $x_\gamma^{obs} < 0.7$ values as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$.

| $P_T^{Jet_1}$ [GeV] | $\langle N_{ch} \rangle$ | $\Delta_{UF}$ | $\Delta_{Sys}$ | $\Delta_{tot}$ |
|---|---|---|---|---|
| 6.9 | 4.38 | 0.04 | 0.12 | 0.12 |
| 10.4 | 5.06 | 0.06 | 0.14 | 0.15 |
| 17.9 | 5.76 | 0.15 | 0.16 | 0.22 |

**Table B.4:** Charged particle average in the **away** region for events with $x_\gamma^{obs} > 0.7$ values as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$.

| $P_T^{Jet_1}$ [GeV] | $\langle N_{ch} \rangle$ | $\Delta_{UF}$ | $\Delta_{Sys}$ | $\Delta_{tot}$ |
|---|---|---|---|---|
| 6.9 | 2.63 | 0.03 | 0.07 | 0.08 |
| 10.4 | 2.57 | 0.05 | 0.07 | 0.09 |
| 17.9 | 2.45 | 0.21 | 0.07 | 0.22 |

**Table B.5:** Charged particle average in the **transverse high activity** region for events with $x_\gamma^{obs} < 0.7$ values as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$.

| $P_T^{Jet_1}$ [GeV] | $\langle N_{ch} \rangle$ | $\Delta_{UF}$ | $\Delta_{Sys}$ | $\Delta_{tot}$ |
|---|---|---|---|---|
| 6.9 | 1.47 | 0.02 | 0.04 | 0.05 |
| 10.4 | 1.47 | 0.02 | 0.04 | 0.05 |
| 17.9 | 1.36 | 0.05 | 0.04 | 0.06 |

**Table B.6:** Charged particle average in the **transverse high activity** region for events with $x_\gamma^{obs} > 0.7$ values as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$.

| $P_T^{Jet_1}$ [GeV] | $\langle N_{ch} \rangle$ | $\Delta_{UF}$ | $\Delta_{Sys}$ | $\Delta_{tot}$ |
|---|---|---|---|---|
| 6.9 | 1.25 | 0.02 | 0.03 | 0.04 |
| 10.4 | 1.18 | 0.03 | 0.03 | 0.04 |
| 17.9 | 1.01 | 0.07 | 0.03 | 0.08 |

**Table B.7:** Charged particle average in the **transverse low activity** region for events with $x_\gamma^{obs} < 0.7$ values as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$.

| $P_T^{Jet_1}$ [GeV] | $\langle N_{ch} \rangle$ | $\Delta_{UF}$ | $\Delta_{Sys}$ | $\Delta_{tot}$ |
|---|---|---|---|---|
| 6.9 | 0.50 | 0.01 | 0.01 | 0.02 |
| 10.4 | 0.47 | 0.01 | 0.01 | 0.02 |
| 17.9 | 0.35 | 0.03 | 0.01 | 0.03 |

**Table B.8:** Charged particle average in the **transverse low activity** region for events with $x_\gamma^{obs} > 0.7$ values as a function of the transverse momentum of the leading jet, $P_T^{Jet_1}$.

| | $\langle N_{ch} \rangle$ | $\Delta_{UF}$ | $\Delta_{Sys}$ | $\Delta_{tot}$ |
|---|---|---|---|---|
| | 1.03 | 0.02 | 0.03 | 0.03 |
| | 0.93 | 0.01 | 0.03 | 0.03 |
| | 0.78 | 0.01 | 0.02 | 0.02 |
| | 0.51 | 0.01 | 0.02 | 0.02 |
| | 0.39 | 0.01 | 0.01 | 0.01 |
| | 0.36 | 0.01 | 0.01 | 0.01 |
| | 0.37 | 0.01 | 0.01 | 0.01 |
| | 0.55 | 0.01 | 0.01 | 0.02 |
| | 0.87 | 0.01 | 0.02 | 0.03 |
| | 1.50 | 0.02 | 0.04 | 0.05 |
| | 1.50 | 0.02 | 0.04 | 0.05 |
| | 0.86 | 0.01 | 0.02 | 0.03 |
| | 0.57 | 0.01 | 0.02 | 0.02 |
| | 0.62 | 0.01 | 0.02 | 0.02 |
| | 0.75 | 0.01 | 0.02 | 0.02 |
| | 0.83 | 0.01 | 0.02 | 0.03 |
| | 0.82 | 0.01 | 0.02 | 0.03 |
| | 0.75 | 0.01 | 0.02 | 0.02 |
| | 0.85 | 0.01 | 0.02 | 0.03 |
| | 0.96 | 0.01 | 0.03 | 0.03 |

**Table B.9:** Charged particle average for events with $x_\gamma^{obs} < 0.7$ values as a function of the azimuthal angle with respect the leading jet, $\Delta\phi^{Jet_1, h\pm}$.

| | $\langle N_{ch} \rangle$ | $\Delta_{UF}$ | $\Delta_{Sys}$ | $\Delta_{tot}$ |
|---|---|---|---|---|
| | 1.13 | 0.02 | 0.04 | 0.04 |
| | 0.81 | 0.01 | 0.02 | 0.03 |
| | 0.51 | 0.01 | 0.01 | 0.02 |
| | 0.24 | 0.01 | 0.01 | 0.01 |
| | 0.15 | 0.01 | 0.00 | 0.01 |
| | 0.13 | 0.01 | 0.00 | 0.01 |
| | 0.17 | 0.01 | 0.01 | 0.01 |
| | 0.35 | 0.01 | 0.01 | 0.01 |
| | 0.68 | 0.01 | 0.02 | 0.02 |
| | 1.47 | 0.02 | 0.04 | 0.05 |
| | 1.47 | 0.02 | 0.04 | 0.05 |
| | 0.69 | 0.01 | 0.02 | 0.02 |
| | 0.36 | 0.01 | 0.01 | 0.01 |
| | 0.36 | 0.00 | 0.01 | 0.01 |
| | 0.40 | 0.00 | 0.01 | 0.02 |
| | 0.45 | 0.00 | 0.01 | 0.02 |
| | 0.49 | 0.00 | 0.02 | 0.02 |
| | 0.47 | 0.01 | 0.01 | 0.02 |
| | 0.64 | 0.01 | 0.02 | 0.02 |
| | 0.97 | 0.02 | 0.03 | 0.03 |

**Table B.10:** Charged particle average for events with $x_{\gamma}^{obs} > 0.7$ values as a function of the azimuthal angle with respect the leading jet, $\Delta\phi^{Jet_1,\ h^{\pm}}$.

# Appendix C

# How to calculate $\sigma(p\bar{p} \to t\bar{t})$ at TeVatron

This appendix is an example how to calculate matrix elements. In particular the cross section $p\bar{p} \to t\bar{t}$ at TeVatron. This are the notes of a lecture given at the Tutorial on QCD and electroweak physics at the LHC: "Collinear Factorization, DGLAP, Jet/Heavy Flavour production, LO/NLO" on Wednesday 18th of October 2006 by Peter Uwer.

TeVatron is a proton anti-proton collider with a center of mass energy of 1.96 TeV. We want to study the top quark production. The basic picture for this process is:

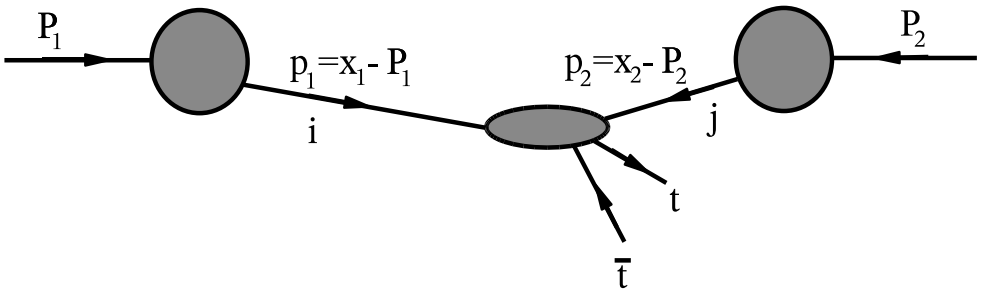


**Figure C.1:** Basic picture for top quark production in a $p\bar{p}$ collision.

From the two hadrons, with 4-momenta $P_1$ and $P_2$ respectively, two partons, $i$ and $j$, enter in the hard interaction, with 4-momentum $p_1$ and $p_2$ respectively. From the hard interaction we produce a $t\bar{t}$ pair.

Important remarks on this picture are:

- Only two partons interact, one from each hadron.

- We will use proton parton density functions, pdf. They provide the probability to have a parton with a momentum fraction between $[x,\, x+dx]$, i.e. a parton $i$ in the hadron $H$ with a momentum fraction $x$ at a given scale $\mu_f$, $F_{i/H}(x,\mu_f)$. They have to be measured experimentally but are universal, i.e. experiment independent.

Thus, from this basic picture and using the factorization theorem we can write the $p\bar{p} \to t\bar{t}$ cross section as:

$$\sigma_{hadr.} \;=\; \sum_{i,j} \int dx_i dx_j \; F_{i/H_1}(x_1,\mu_f)\; F_{j/H_2}(x_2,\mu_f)\; d\sigma(i(x_1P_1), j(x_2P_2) \to t\bar{t})$$

where $x_1$ and $x_2$ are the momentum fractions of the partons $i$ and $j$ in the hadrons $H_1$ and $H_2$ with momentum $P_1$ and $P_2$ respectively. The structure functions $F_{i/H_1}(x_1,\mu_f)$ and $F_{j/H_2}(x_2,\mu_f)$ are measured experimentally and can be taken from other experiments, as for example HERA. Now we have to calculate $d\sigma(ij \to t\bar{t})$. This can be written as follows:

$$\begin{aligned} \sigma(ij \to t\bar{t}) \;&=\; \mathcal{N}\frac{1}{2s} dLips\, |\mathcal{T}(ij \to t\bar{t})|^2 \;\; where \\ & dLips = (2\pi)^4 \delta(P_1 + P_2 - \sum_k p_k) \prod_k \frac{d^3\vec{k}_k}{(2\pi)^3 2E_k} \end{aligned}$$

where $\mathcal{T}$ is the transition matrix and $k$ runs over all the final particles, in our case $t$ and $\bar{t}$. We want to remind you that dLips is a Lorentz invariant quantity and to point that it also ensures 4-momentum conservation.

In this last equation we have isolated the transition part from the initial state $|P_1P_2\rangle$ to the final state $|p_1p_2\rangle$. This is done in the term: $\mathcal{T}(ij \to t\bar{t})$. Note that we are using the $S - matrix$ formalism where $S = 1 + i\mathcal{T}$ and this is precisely where the specific theory for a given process enters. The Lagrangian QCD:

$$\begin{aligned} \mathcal{L}_{QCD} \;&= \textstyle\sum_i \bar{\Psi}_i (i \not{D} - m_i)\Psi_i - \frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} \qquad where \;\; \not{D} = \gamma^\mu D_\mu \; and \\ & D_\mu = \partial_\mu - igT_a A^a_\mu \\ & G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + \underbrace{g f_{abc} A^b_\mu A^c_\nu}_{non-abelian\;\; component} \end{aligned}$$

From this Lagrangian we extract the Feynman rules, which will help us to calculate the $\mathcal{T}(ij \to t\bar{t})$ term. At leading order, LO, there are several diagrams contributing to the $t\bar{t}$ production. These are shown in figure C.2.

These diagrams consist of two contributions for $t\bar{t}$ production: $q\bar{q} \to t\bar{t}$ and $gg \to t\bar{t}$. We will calculate the first one and will comment the second in the next sections.

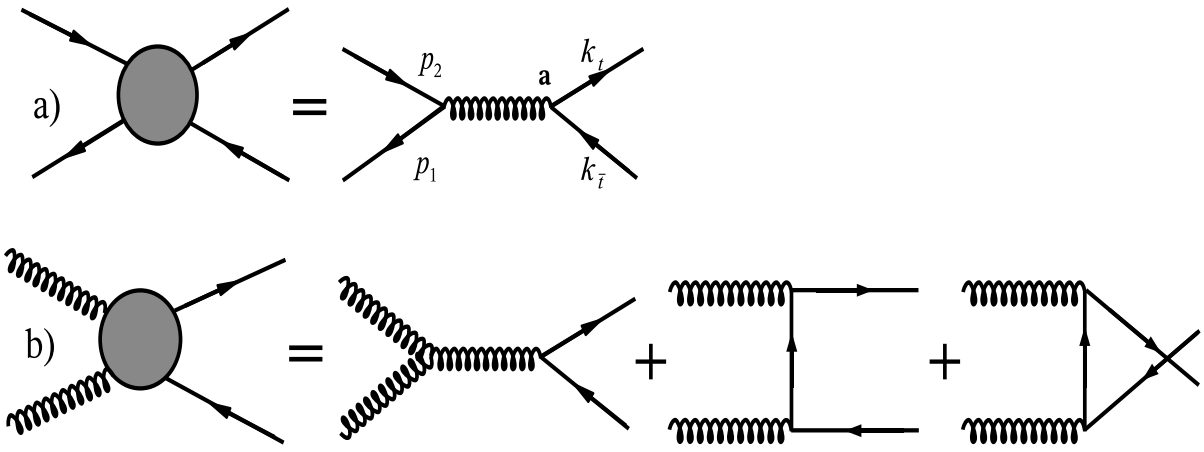


**Figure C.2:** Leading order diagrams for $t\bar{t}$ production at TeVatron.

## C.1 $q\bar{q} \to t\bar{t}$

The kinematics for this process is depicted in figure C.2 a). One quark from one hadron and one anti-quark from the other hadron, with 4-momentum $p_1$ and $p_2$ respectively interact and produce a top and an anti-top quark pair, with 4-momenta $k_t$ and $k_{\bar{t}}$.

Using the Feynman rules we can calculate the matrix element:

$$\begin{aligned} i\mathcal{T} &= \bar{u}(k_t)\ (-i)g_s\gamma_\mu T^a\ v(k_{\bar{t}})\tfrac{-ig^{\mu\nu}\bar{v}()\delta^{ab}}{(k_t+k_{\bar{t}})^2+i\epsilon}\bar{v}(p_2)(-i)g_s\gamma^\nu T^b u(p_1) \\ &= ig_s^2\bar{u}(k_t)\gamma_\nu T^a v(k_{\bar{t}})\bar{v}(p_2)\gamma^\nu T^a u(p_1)\tfrac{1}{s+i\epsilon} \end{aligned}$$

To calculate the cross section we need $|\mathcal{T}|^2$:

$$\begin{aligned} |\mathcal{T}|^2 &= g_s^4\tfrac{1}{s^2}\ \bar{u}(k_t)\ \gamma_\nu T^a\ v(k_{\bar{t}})\ \bar{v}(p_2)\ \gamma^\nu T^a\ u(p_1) \\ &\quad \bar{v}(k_{\bar{t}})\ \gamma_\mu T^b\ u(k_t)\ \bar{u}(p_1)\ \gamma^\mu T^b\ v(p_2) \end{aligned} \tag{C.1}$$

Since we do not have any information about the colour of the incoming and outgoing partons we sum and average over colour:

$$\begin{aligned} \frac{1}{3\cdot 3}\sum_{t,\bar{t},q,\bar{q}} (T^a)_{t\bar{t}}\left(T^b\right)_{t\bar{t}}(T^a)_{q\bar{q}}\left(T^b\right)_{q\bar{q}} &= \tfrac{1}{9}\ Tr[T^a\ T^b]\ Tr[T^a\ T^b] \\ &= \tfrac{1}{9}\ \tfrac{1}{2}\ \delta^{ab}\ \tfrac{1}{2}\ \delta^{ab} = \tfrac{1}{9}\ \tfrac{1}{4}\ \underbrace{\delta^{aa}}_{8} \\ in\ SU(N) \Longrightarrow \qquad &= \tfrac{1}{N^2}\ \tfrac{1}{4}\ (N^2-1) \end{aligned} \tag{C.2}$$

To calculate further equation C.1 we need to use:

$$\sum_s u_\alpha(k_t, s)\ \bar{u}_\beta(k_t, s) = (\not{k}_t\ +\ m_t)_{\alpha\beta}$$
$$\sum_s v_\alpha(k_{\bar{t}}, s)\ \bar{v}_\beta(k_{\bar{t}}, s) = (\not{k}_{\bar{t}}\ -\ m_t)_{\alpha\beta} \tag{C.3}$$

plugging all together and summing and averaging over spin:

$$\sum_{spin} |\mathcal{T}|^2\ = c\ g_s^4\ \tfrac{1}{s^2}\ Tr[(\not{k}_t\ +\ m)\ \gamma_\nu\ (\not{k}_{\bar{t}}\ - m)\ \gamma_\mu]\ Tr[\not{p}_2\ \gamma^\nu\ \ \not{p}_1\ \gamma^\mu] \tag{C.4}$$
$$= c\ g_s^4\ \tfrac{1}{s^2}\ 4\ (2\ + (z^2-1)\ \beta^2)\ s^2 \tag{C.5}$$

where $z = cos\theta$ calculated in the center of mass, $dLips = \frac{1}{16\pi}\beta dz$ and we used $\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$ to simplify the expressions.

Finally:

$$d\sigma\ = \underbrace{\frac{1}{2\cdot 2}}_{spin}\ \underbrace{\frac{1}{3\cdot 3}}_{colour}\ \tfrac{1}{2s} 2 g_s^4 \tfrac{1}{s^2} 4 (2\ + (z^2-1)\ \beta^2)\ s^2 \tfrac{1}{16\pi}\beta dz \tag{C.6}$$
$$= \tfrac{1}{9}\ \pi\alpha^2\ \tfrac{1}{s}\ \beta(2 + (z^2-1)\beta^2)\ dz \tag{C.7}$$

where $\alpha = \frac{g_s^2}{4\pi}$

## C.2 $gg \to t\bar{t}$

Although we are not going to calculate this process, depicted in figure C.2 b), we want to do some remarks:

- Since there are two gluons, which carry colour, colour is more involved.
- For the polarization of the gluon we use:

$$\sum_\lambda \epsilon_\lambda^\mu(p_1)\epsilon_\lambda^{*\nu}(p_2)\ \ =\ \ -g^{\mu\nu} + \frac{n^\mu p_1^\nu + n^\nu p_1^\mu}{n\ p}\ \ where\ n^2 = 0 \tag{C.8}$$

## C.3 Hadronic cross section.

Back to the beginning, we wrote the hadronic cross section as:

$$d\sigma_{hadr.} = \sum_{i,j} \int dx_i dx_j \; F_{i/H_1}(x_1, \mu_f) \; F_{j/H_2}(x_2, \mu_f) \; d\sigma(s_{tevatron} x_1^i x_2^j, z)$$

where we want to remind that: $s = (p_1 + p_2)^2 = 2p_1 \; p_2$ and therefore: $s = 2x_1 P_1 \; x_2 P_2 = x_1 x_2 s_{tevatron}$. We want to point also that $F_{i/H}$ and $\alpha_s$ depend on a scale $\mu$. However, physical predictions cannot depend on a scale. This means:

- $F_{i/H}$ and $\alpha_s$ are not direct observables and are defined only in an specific approach.
- Since the predictions are independent on the scale $\mu$, we can use in principle any value we like.
- In practice, we choose a "good" value of $\mu$, so that perturbation theory converges well.

Beyond LO we find terms like: $ln\left(\frac{\mu^2}{s_{ij}}\right)$, where $s_{ij}$ is some typical energy scale of the process. This is a motivation to choose $\mu$ of the order of the typical energy scale of the problem, so that the next correction terms are small. We choose $\mu_f = \mu_r = 2m_t$, so that $F_{i/H_1}(x_1, \mu_f) = F_{i/H_1}(x_1, 2m_t)$. With this choice: $\alpha_s(2m_t) \approx 0.1$

# Appendix D

# PYTHIA parameters

In this appendix the different parameters in PYTHIA used in this analysis are given.

| Parameter | Value | Short description |
|---|---|---|
| CKIN (3) | 2.5 GeV | Allowed $\hat{p}_t$ values for hard $2 \rightarrow 2$ processes. |
| MSTP (51) | 10041 | Choice of proton parton-distribution set: CTEQ6l (LO fit/NLO $\alpha_s$). |
| MSTP (52) | 2 | Proton parton-distribution set from PDFLIB. |
| MSTP (55) | 395 | Choice of photon parton-distribution set: SaS-G 1D (ver.2) LO. |
| MSTP (56) | 2 | Photon parton-distribution set from PDFLIB. |
| MSTP (93) | 5 | Primordial $k_t$ distribution in photon (power-like: $dk_t^2/(1GeV + k_t^2)$ and upper cut-off at 5 GeV). |

**Table D.1:** Parameters common to all simulations done with the PYTHIA Monte Carlo generator.

| Parameter | Short description |
|---|---|
| MSTP (81) | Master switch for multiple parton interactions: 0 = off, 1 = on. |

**Table D.2:** Parameter used to switch on/off the multiple parton interaction model.

| Parameter | Short description |
|---|---|
| MSTP (61) | Master switch for initial-state radiation: 0 = off, 1 = on. |
| MSTP (71) | Master switch for final-state radiation: 0 = off, 1 = on. |

**Table D.3:** Parameters used to switch on/off initial or final state radiation in section 2.4.

| Parameter | Value (variation) | Short description |
|---|---|---|
| PARP (85) | 0.33 (0.33-1) | Probability that additional interactions give two gluons with colour connexions to the 'nearest neighbours'. |
| PARP (86) | 0.66 (0.66-1) | Probability that additional interactions give two gluons. The remaining fraction is supposed to be quark-antiquark pairs. |

**Table D.4:** Parameters used to vary the colour string length in section 6.2.3.

| Parameter | Value (variation) | Short description |
|---|---|---|
| PARP (67) | 4 (1-4) | The $Q^2$ of the hard scattering is multiplied by this factor to define the maximum virtuality allowed in space-like showers. |

**Table D.5:** Parameter used to vary the upper scale for the parton showers in section 6.2.4.

| Parameter | Value (variation) | Short description |
|---|---|---|
| PARP (81) | 1.9 (1.5-1.9 GeV) | Effective minimum transverse momentum $p_{\perp,min}$ for multiple interactions. |

**Table D.6:** Parameter used to vary the average number of MPI per event, section 6.3.

# Study of Multiple Interactions in photoproduction at HERA

ICHEP Conference Note

30th July - 5 August 2008, Philadelphia.

**H1prelim-08-036** Submitted to
**34th International Conference on High Energy Physics, ICHEP2008**, 30th July - 5th August, Philadelphia
Abstract: **xx-xxx**
Parallel Session **QCD/Lattice**
*Electronic Access: www-h1.desy.de/h1/www/publications/conf/conf_list.html*

# Study of Multiple Interactions in photoproduction at HERA (version of July 22, 2008)

H1 Collaboration

**Abstract**

Photoproduction data of HERA-I are analysed by requiring dijets with transverse momenta of at least 5 GeV. The two jets define in azimuth a towards region (leading jet) and an away region (2nd jet) and transverse regions between them. The charged particle multiplicity is measured in these regions as a function of the variables $x_\gamma$ and $P_T$ (leading jet). PYTHIA predictions including contributions of multi-parton interactions are able to describe the measurement, whereas without them they lie below the measurements, especially at low $x_\gamma$, the region of enhanced contributions from the resolved photon.

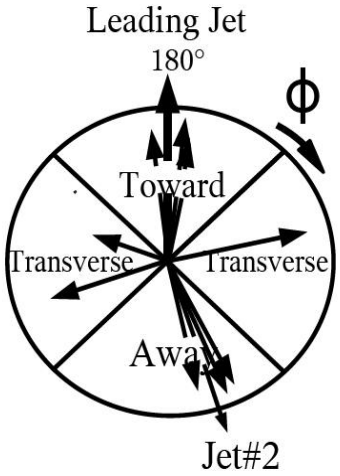


Figure 1: *Toward, away and transverse regions defined. In the toward region the leading jet is always at $\phi = 180^o$ and is the reference to define the other regions.*

# 1 Introduction

In quasi-real photoproduction ($Q^2 \sim 0$) the photon has a point-like as well as a hadronic (resolved) component. Measurements in photoproduction at HERA have the advantage that the transition from a point-like photon towards a resolved photon can be studied in detail as function of the variable $x_\gamma$. Multi-parton interactions are expected within the model of [1, 2] for resolved photons ($x_\gamma < 1$) but not for the point-like photons which have $x_\gamma \sim 1$. This fact was already used for multi-parton studies at HERA in [3], where it was shown that jet production and energy flow are well described when including multi-parton interactions.

In the photoproduction of dijets at HERA the underlying event can be studied in a fashion similar to the studies done at the TeVatron [4]. The underlying event is here defined as everything in addition to the lowest order process, here dijet production. The underlying event will consist of contributions coming from higher order QCD radiation, simulated by parton showers, hadronisation and also of multi-parton interactions. This multi-parton interactions take place when the density of partons in the colliding beams is large enough that more than one interaction happens within one collision.

Multi-parton interactions have been required to describe the transverse momentum and particle multiplicity distributions in a region transverse to the jets at the TeVatron, see Fig.1. Here an analysis is presented which follows closely this analysis to investigate the effect of the underlying event and multi-parton interactions in photoproduction at HERA.

# 2 Monte Carlo Model

The full hadronic final state is simulated with the PYTHIA Monte Carlo event generator including the lowest order matrix element processes for direct and resolved photoproduction together

| Description | Used | Default |
|---|---|---|
| MSTP(93) | 5 | 1 |
| PARP(82) | 1.2 | 2 |
| PARP(99) | 0.6 | 1 |
| PARP(85) | 0.33 (1.) | 0.9 |
| PARP(86) | 0.66 (1.) | 0.95 |
| PARP(83) | 0.5 | 0.5 |
| PARP(84) | 0.4 | 0.4 |
| PARP(89) | 1800 | 1800 |
| PARP(90) | 0.25 | 0.16 |

Table 1: The relevant parameters for the PYTHIA simulation.

with parton showers to simulate higher order QCD corrections and hadronisation. In PYTHIA the Lund String fragmentation is implemented.

PYTHIA includes multi-parton interactions (MPI): several partons of the proton can interact with several partons of the resolved photon. The details of the multi-parton interaction model is described in [1,2]. For the photoproduction case only a simple version of this model is available. A similar model was used also to describe the TeVatron measurements.

When multi-parton interactions occur, the formation of colour neutral strings or clusters is far from trivial as colour exchange might happen between the different interactions. In the Lund String model long strings will produce a larger multiplicity of hadrons compared to short strings. The transverse energy, however, is mainly driven by the high-$p_t$ hard scattering process. Thus the particle multiplicity together with the average transverse momentum distribution yields non-trivial information about the colour structure of jet events.

Events are generated with the PYTHIA Monte Carlo event generator with multi-parton interactions switched on. The proton pdf was CTEQ6L [5] and the photon pdf was SaS1D [6]. All generated events are passed through the full simulation of the H1 apparatus and are reconstructed using the same program chain as for the data.

The final results of the measurement are then compared with predictions from PYTHIA at hadron level. The different settings are given in Tab.1. It was explicitly checked that same set of parameters in PYTHIA 6.2 and PYTHIA 6.4 give the same results. The measurements are compared with Monte Carlo simulations with and without multi-parton interactions and with different options for the colour configuration, as given by the parameter PARP(85), PARP(86) for short (PARP(85)=PARP(86)=1) and long string (PARP(85)=0.33, PARP(86)=0.66) configurations.

# Event Selection

Data are recorded in the running period 1999-2000, $\mathcal{L} = 48\ pb^{-1}$, where HERA was operated with positron and proton beams of $E_e = 26.7$ GeV and $E_p = 920$ GeV respectively giving a center-of-mass energy of $\sqrt{s} = 318$ GeV.

Photoproduction events are selected by requiring a signal in the electron tagger which has acceptance in the inelasticity range $0.3 < y < 0.65$, which results in $Q^2 < 0.01$ GeV$^2$. The reconstructed primary vertex position, $z-$vertex, has to be within 35 cm of the nominal one.

Jets are selected with the inclusive $k_t$-jet algorithm using as input combined objects, i.e. four-vectors objects, where information from track and/or cluster is combined. At least two jets are required with $P_T^{jet} > 5$ GeV within a pseudorapidity range in the laboratory frame of $-1.5 < \eta_{lab} < 1.5$. The jet with the highest $P_T^{jet}$ defines the leading jet, $Jet_1$. From the jets the value $x_\gamma$ is calculated:

$$x_\gamma = \frac{\sum_{h \in Jet_1}(E - P_z) + \sum_{h \in Jet_2}(E - P_z)}{\sum_h (E - P_z)} \tag{1}$$

Tracks are selected having a $P_T^{track} > 150$ MeV within $|\eta| < 1.5$. The difference in azimuthal angle $\Delta\phi = \phi_{Jet_1} - \phi_i$ with $i$ running over all charged particles satisfying the above cuts, is calculated, with the leading jet set to $\phi = 180^o$.

Four different regions in azimuth with respect to the leading jet are defined, see Fig.1:

- the toward region (including the leading jet) with $120^o < \Delta\phi < 240^o$
- the away region with $300^o < \Delta\phi < 60^o$
- the transverse region with $60^o < \Delta\phi < 120^o$ and $240^o < \Delta\phi < 300^o$

In the transverse region the scalar sum of the transverse momenta $E_t^{sum} = \sum_i^{tracks} P_T^i$ is calculated and the region with the higher $E_t^{sum}$ is defined as the *high activity* region. The direction of $\phi$ is defined such, that the *high activity* is always $240^o < \Delta\phi < 300^o$.

The average multiplicity distributions are measured as a function of the angular difference $\Delta\phi$ and as a function of the leading jet transverse momentum $P_T^{jet}$ for the four different regions in $\Delta\phi$. The distributions are corrected to charged stable hadrons level using an iterative Bayes unfolding method (see [7]).

## Systematic Uncertainties

In this section, the systematic uncertainties associated with the measurement of photoproduction of dijets are discussed. The uncertainties on the signal extraction are determined by varying the following quantities:

- $\pm 2$ % for the track finding in CJC, $\pm 1$ % for the vertex reconstruction using central tracks and $\pm 1.5$ % for uncertainties in the simulation of nuclear interaction resulting in a $\pm 2.7$ % independent and uncorrelated uncertainty independent of $P_T$ and $\eta$ of the charged tracks.
- $\pm 2$ % of the hadronic energy scale of the calorimeter for the construction of combined objects.

- In case the difference in $E_t^{sum}$ between the high and low activity transverse region is less than $160$ MeV the probability to misidentify the high and low activity regions is estimated to be $10\%$.

- $2-4$ % uncertainty from the limited statistics used in the unfolding.

The individual effects of the above experimental uncertainties are combined in quadrature, yielding a total uncertainty of $\sim 3\%$ on the signal expectation. The largest contributions to this uncertainty arise from systematics attributed to tracking ($2.7\%$ nearly constant over all regions and phase space) and to the hadronic energy scale ( $0.5\%$ in the toward and away regions and $0.5\% - 1.5\%$ in the transverse regions).

# 3 Results

The results shown here are always shown for two regions in $x_\gamma$: a resolved photon enriched region with $x_\gamma < 0.7$ and a point-like photon enriched region with $x_\gamma > 0.7$. The average track multiplicity as a function of $\Delta\phi$ is shown in Fig. 2 for the two regions in $x_\gamma$. One can see clearly the leading jet which is by definition at $\Delta\phi \sim 180^o$, the second jet at $\Delta\phi \sim 0^o$ and $\Delta\phi \sim 360^o$ as well as contributions of a third jet which is in the region of $\Delta\phi > 240^o$. The data are compared to predictions of PYTHIA. In the large $x_\gamma$ region the effect of multi-parton interaction is very small, in the region of $x_\gamma < 0.7$ the inclusion of multi-parton interaction clearly improves the description in the transverse regions as well as in the toward and away regions where it contributes as a pedestal to the track multiplicity.

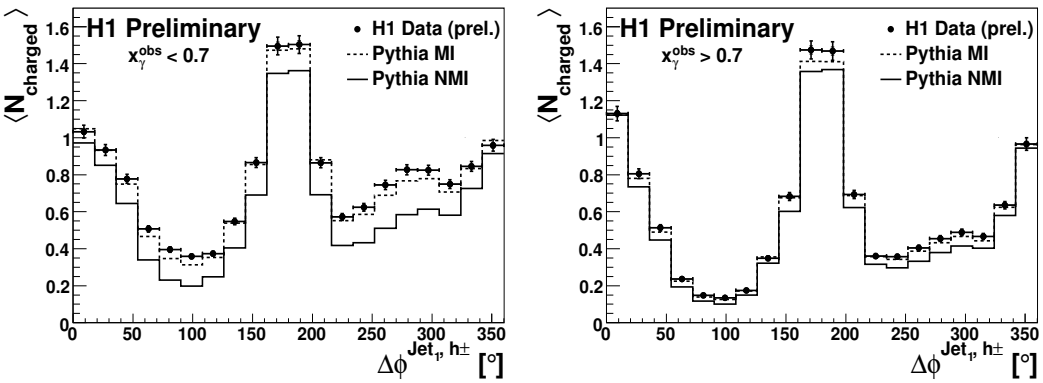


Figure 2: Charged particle multiplicity for $x_\gamma < 0.7$ (left) and for $x_\gamma > 0.7$ (right). The leading jet axis is by definition at $180^o$. Data is compared to PYTHIA with and without MPI.

In Fig. 3 the average charge particle multiplicity is shown for the two regions in $x_\gamma$ for the toward and away regions. In general the average track multiplicity rises with increasing $P_T^{Jet_1}$ from around 4-5 particles at $P_T^{Jet_1} \sim 5$ GeV to around 7 particles $P_T^{Jet_1} > 15$ GeV also depending on $x_\gamma$. In the region of $x_\gamma > 0.7$ the measurements are reasonably well described with

a simulation containing only one hard interaction together with parton showers and hadronisation, whereas in the region $x_\gamma < 0.7$ this is clearly not enough, especially at the lower values of $P_T^{Jet_1}$. Including multi-parton interactions brings the prediction in good agreement with the measurement.

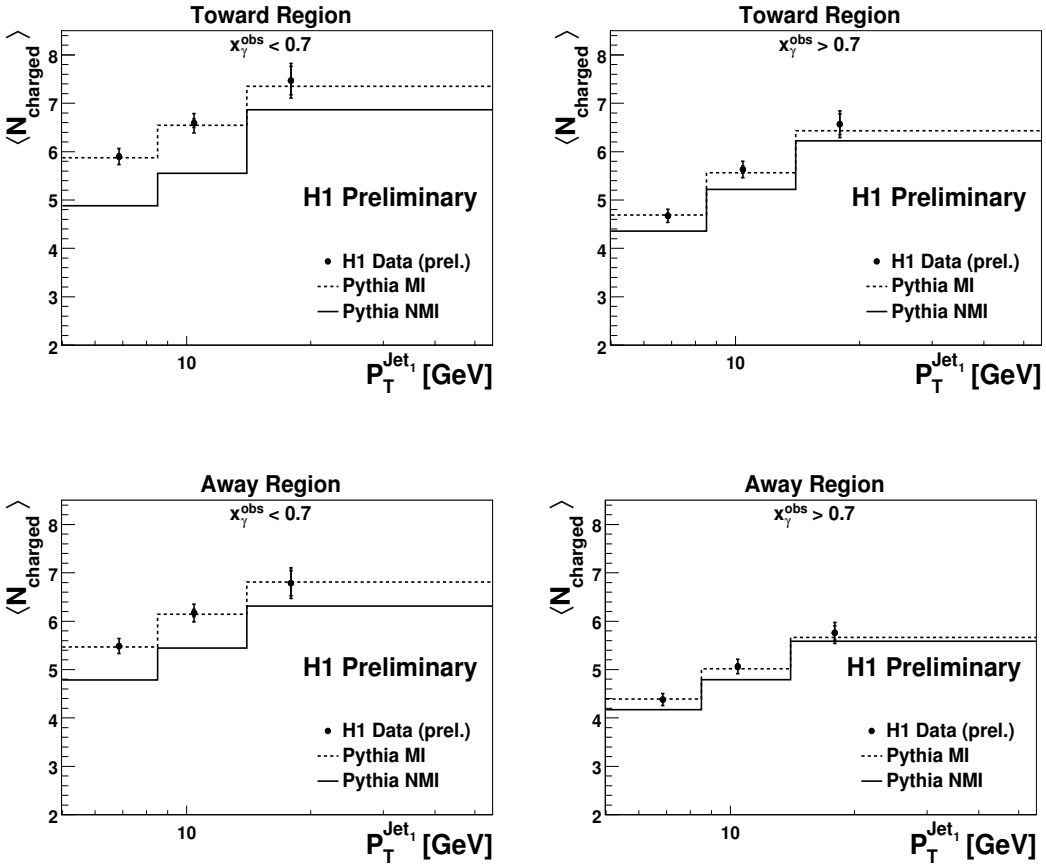


Figure 3: Charged particle multiplicity for $x_\gamma < 0.7$ (left) and for $x_\gamma > 0.7$ (right). The toward regions, upper plots, contain the leading jet whereas the second jet is usually found in the away regions, down. Data is compared to PYTHIA with and without MPI.

In Fig. 4 the average charged particle multiplicity in the traverse region is shown for the two regions in $x_\gamma$. Here the average charged particle multiplicity ranges between 0.5 to 2.5 particles depending on $x_\gamma$ and $P_T^{Jet_1}$. In the region of $x_\gamma > 0.7$ the measurements are again reasonably well described without including MPI, whereas the region of $x_\gamma < 0.7$ requires a larger average multiplicity than predicted from a simulation without MPI. The simulation including MPI gives a reasonable description of the measurement over the full phase space region.

It is interesting to investigate whether the measurements can also be described by including more higher order QCD radiation, which can be approximately simulated by changing the upper scale in the initial and final state parton shower simulation, PS. Setting the upper scale to four times the scale of the hard scattering process results in harder partons from the parton shower.

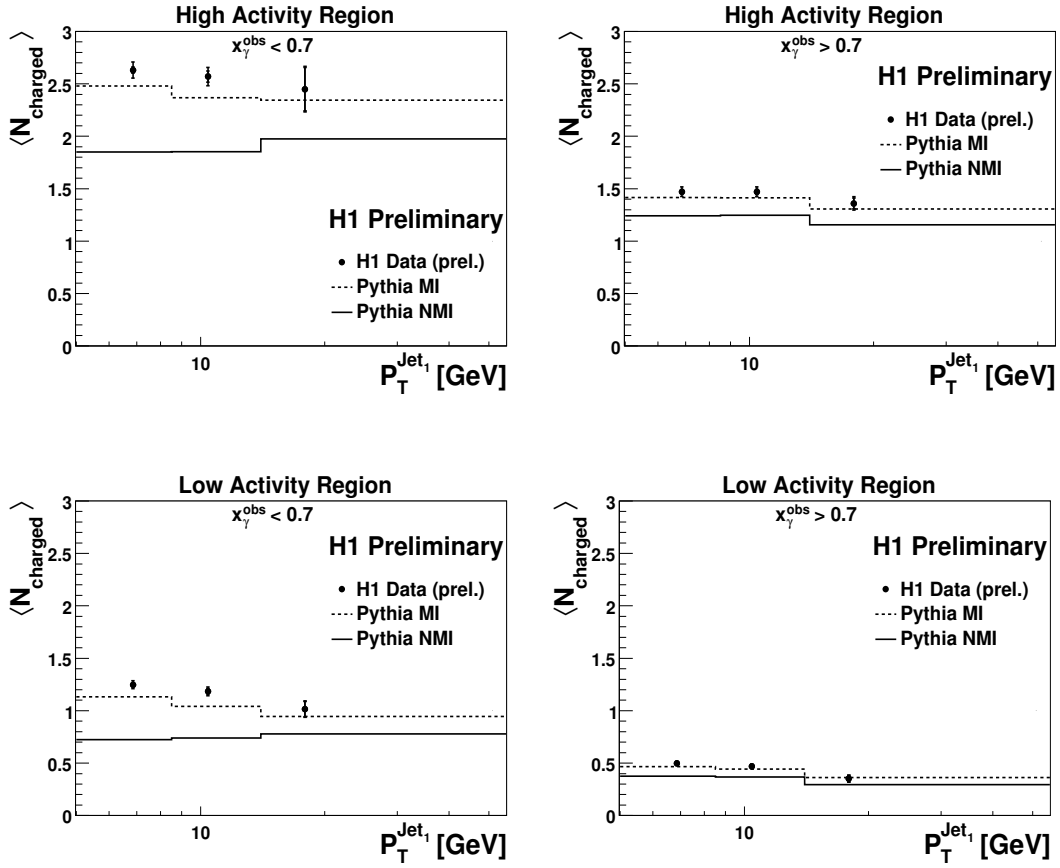


Figure 4: Charged particle multiplicity for $x_\gamma < 0.7$ (left) and for $x_\gamma > 0.7$ (right). The transverse high activity regions, upper plots, are defined as the transverse region with a higher $E_t^{sum}$ compared to the low activity regions, down. Data is compared to PYTHIA with and without MPI.

The comparison of a simulation with an upper scale four times the hard scattering scale and a simulation with this upper scale set to the hard scattering is shown in Fig. 5. Only in the high activity region a effect of the parton shower can be seen, but this is not enough to bring the simulation close to the measurement, indicating that higher order contributions cannot account for the high average particle multiplicity in the transverse regions.

Since the particle multiplicity is sensitive to the details of the colour structure of the final state, the effect of chosing short or long string configurations is shown in Fig. 6. In the present simulation the long string configuration is preferred.

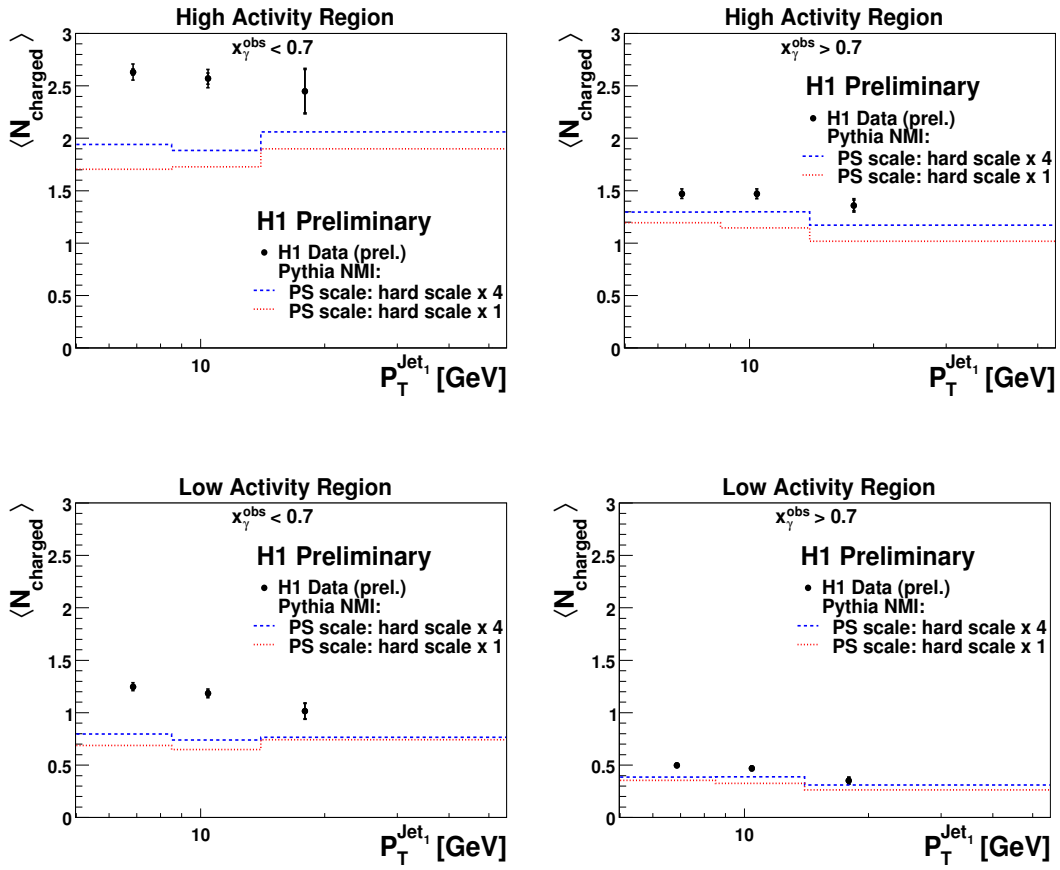


Figure 5: Charged particle multiplicity for $x_\gamma < 0.7$ (left) and for $x_\gamma > 0.7$ (right). The transverse high activity regions, upper plots, are defined as the transverse region with a higher $E_t^{sum}$ compared to the low activity regions, down. The comparison of PYTHIA without MPI with an upper scale for the PS four times the hard scattering scale and with this scale set to the hard scattering is shown

# 4 Conclusion

The average charged particle multiplicity in photoproduction of dijet events has been measured as a function of the azimuthal angle with respect to the leading jet and compared to PYTHIA. The charged particle multiplicity has been investigated for two different regions in $x_\gamma$ separating the point-like from the hadronic (resolved) photon and as a function of the leading jet transverse momentum. In the region of $x_\gamma > 0.7$ the charged particle multiplicity is reasonably well described by PYTHIA including one hard scattering process supplemented with parton showers and hadronisation but without MPI. The region of $x_\gamma < 0.7$ shows clearly, that the charged particle multiplicity is larger than what is expected from this simulation, and is well described when multi-parton interactions are included. The multi-parton interaction simulation describes the measurement better when the colour structure of the final state is chosen such that long

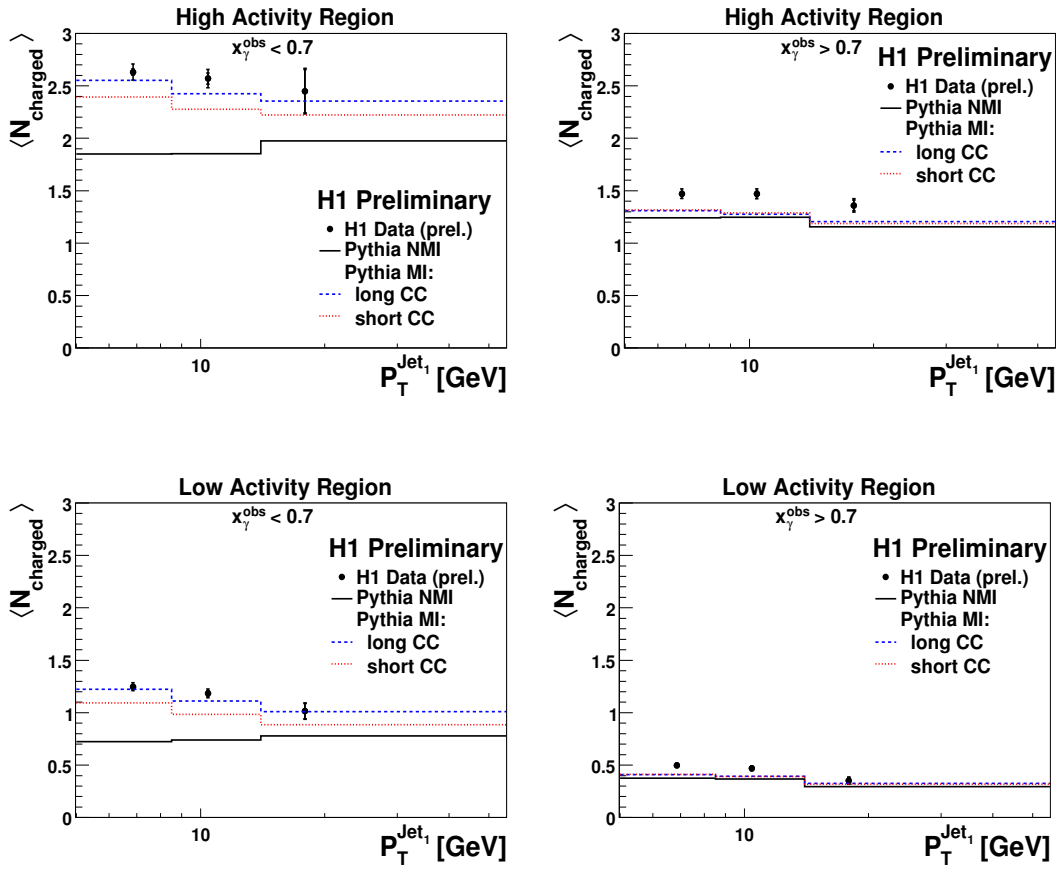


Figure 6: Charged particle multiplicity for $x_\gamma < 0.7$, left, and for $x_\gamma > 0.7$, right. The transverse high activity regions, upper plots, are defined as the transverse region with a higher $E_t^{sum}$ compared to the low activity regions, down. Here data is compared to PYTHIA without MPI and PYTHIA with MPI and high probability for long colour string connections, long CC, and short colour string connections, short CC.

string lengths are preferred.

It was also shown, that the charged particle multiplicity in the transverse region cannot be accounted for by increasing the amount of initial and final state radiation.

In the region where the photon develops a hadronic structure the measurement supports the picture where multi-parton interactions occur, very similar to what has been observed at the TeVatron.

# Acknowledgments

We are grateful to the HERA machine group whose outstanding efforts have made this experiment possible. We thank the engineers and technicians for their work in constructing and

maintaining the H1 detector, our funding agencies for financial support, the DESY technical staff for continual assistance and the DESY directorate for support and for the hospitality which they extend to the non DESY members of the collaboration.

# 5 References

# References

[1] T. Sjöstrand, L. Lonnblad and S. Mrenna, "PYTHIA 6.2: Physics and manual", (2001), hep-ph/0108264.

T. Sjöstrand, S. Mrenna and P. Skands, "PYTHIA 6.4 Physics and Manual", JHEP 0605:026,2006. (2006), hep-ph/0603175.

[2] T. Sjöstrand and M. van Zijl, "A multiple-interaction model for the event structure in hadron collisions", Phys. Rev. D36 (1987), 7, 2019–2041.

[3] S. Aid *et al.* [H1 Collaboration], "Jets and Energy Flow in Photon-Proton Collisions at HERA", Z. Phys. C 70 (1996) 17.

[4] T. Affolder, *et al.* [CDF Collaboration], "Charged jet evolution and the underlying event in proton anti-proton collisions at 1.8-TeV", Phys. Rev. D65 (2002) 092002.

[5] J. Pumplin, *et al.*, "New generation of parton distributions with uncertainties from global QCD analysis", JHEP **0207** (2002) 012 [arXiv:hep-ph/0201195].

[6] G. A. Schuler and T. Sjostrand, "Parton Distributions of the Virtual Photon", Phys. Lett. B **376** (1996) 193 [arXiv:hep-ph/9601282].

[7] D'Agostini, G., "A Multidimensional unfolding method based on Bayes' theorem", Nucl. Instrum. Meth. A362 (1995), 487-498.

# List of Figures

# List of Tables

# References


[1] J. D. Bjorken and E. A. Paschos, "Inelastic Electron Proton and gamma Proton Scattering, and the Structure of the Nucleon," *Phys. Rev.*, vol. 185, pp. 1975–1982, 1969. 5

[2] R. Devenish and A. Cooper-Sarkar, *"Deep Inelastic Scattering"*, ch. 3. Oxford University Press, 2004. 9

[3] R. Brock *et al.*, "Handbook of perturbative QCD; Version 1.1: September 1994," FERMILAB-PUB-94-316. 9

[4] D. E. S. J.C. Collins and G. Sterman, *"Factorization of Hard Processes in QCD"*. Pertubative QCD, Ed.:A.H.Mueller (World Scientific), 1989. 11

[5] B. Andersson, *"The Lund Model"*. Cambridge University Press, 1998. 11

[6] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, "Parton Fragmentation and String Dynamics," *Phys. Rept.*, vol. 97, pp. 31–145, 1983. 11, 12

[7] T. Sjostrand, "High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4," *Comput. Phys. Commun.*, vol. 82, pp. 74–90, 1994. 11

[8] B. R. Webber, "A QCD Model for Jet Fragmentation Including Soft Gluon Interference," *Nucl. Phys.*, vol. B238, p. 492, 1984. 11

[9] http://webber.home.cern.ch/webber/*hw*65_*manual.html*. 11, 24

[10] B. R. Martin, "Particle physics," Chichester, UK: Wiley (1997) 366 p. (The Manchester physics series). 12

[11] M. Erdmann, "The partonic structure of the photon: Photoproduction at the lepton proton collider HERA," *Springer Tracts Mod. Phys.*, vol. 138, pp. 1–114, 1997. 13

[12] C. Friberg and T. Sjostrand, "Jet production by real and virtual photons," 1999. 13

[13] C. F. von Weizsacker, "Radiation emitted in collisions of very fast electrons," *Z. Phys.*, vol. 88, pp. 612–625, 1934. 15

[14] E. J. Williams, "Nature of the high energy particles of penetrating radiation and status of ionization and radiation formulae," *Phys. Rev.*, vol. 45, pp. 729–730, May 1934. 15

[15] G. Grindhammer, "Photon structure," 1998. 15

[16] G. A. Schuler and T. Sjostrand, "Parton Distributions of the Virtual Photon," *Phys. Lett.*, vol. B376, pp. 193–200, 1996. 15, 34

[17] M. Drees and R. M. Godbole, "Virtual photon structure functions and the parton content of the electron," *Phys. Rev.*, vol. D50, pp. 3124–3133, 1994. 15

[18] L. E. Gordon and J. K. Storrow, "The Single jet inclusive cross-section at HERA in next-to- leading order QCD," *Phys. Lett.*, vol. B291, pp. 320–324, 1992. 17

[19] T. Sjostrand, "A Model for Initial State Parton Showers," *Phys. Lett.*, vol. B157, p. 321, 1985. 18

[20] M. Bengtsson and T. Sjostrand, "PARTON SHOWERS IN LEPTOPRODUCTION EVENTS," *Z. Phys.*, vol. C37, p. 465, 1988. 18

[21] M. Ciafaloni, "Coherence Effects in Initial Jets at Small q**2 / s," *Nucl. Phys.*, vol. B296, p. 49, 1988. 19

[22] S. Catani, F. Fiorani, and G. Marchesini, "QCD Coherence in Initial State Radiation," *Phys. Lett.*, vol. B234, p. 339, 1990. 19

[23] S. Catani, F. Fiorani, and G. Marchesini, "Small x Behavior of Initial State Radiation in Perturbative QCD," *Nucl. Phys.*, vol. B336, p. 18, 1990. 19

[24] G. Marchesini, "QCD coherence in the structure function and associated distributions at small x," *Nucl. Phys.*, vol. B445, pp. 49–80, 1995. 19

[25] R. K. Ellis, W. J. Stirling, and B. R. Webber, "QCD and collider physics," *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.*, vol. 8, pp. 1–435, 1996. 19

[26] S. Catani, M. Ciafaloni, and F. Hautmann, "High-energy factorization and small x heavy flavor production," *Nucl. Phys.*, vol. B366, pp. 135–188, 1991. 20

[27] J. C. Collins and R. K. Ellis, "Heavy quark production in very high-energy hadron collisions," *Nucl. Phys.*, vol. B360, pp. 3–30, 1991. 20

[28] E. M. Levin, M. G. Ryskin, Y. M. Shabelski, and A. G. Shuvaev, "Heavy quark production in semihard nucleon interactions," *Sov. J. Nucl. Phys.*, vol. 53, p. 657, 1991. 20

[29] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, "Longitudinally invariant K(t) clustering algorithms for hadron hadron collisions," *Nucl. Phys.*, vol. B406, pp. 187–224, 1993. 21, 22

[30] http://home.thep.lu.se/ torbjorn/pythiaaux/recent.html. 24

[31] http://projects.hepforge.org/cascade/. 24

[32] H. Jung, "CCFM prediction on forward jets and F2: Parton level predictions and a new hadron level Monte Carlo generator CASCADE," 1998. 24

[33] H. Jung and G. P. Salam, "Hadronic final state predictions from CCFM: The hadron-level Monte Carlo generator CASCADE," *Eur. Phys. J.*, vol. C19, pp. 351–360, 2001. 24

[34] F. Abe *et al.*, "Measurement of double parton scattering in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV," *Phys. Rev. Lett.*, vol. 79, pp. 584–589, 1997. 28, 140

[35] J. Gaudean. Ph.D. Dissertation, Universiteit Antwerpen, 1984. 28, 29

[36] S. Aid *et al.*, "Jets and Energy Flow in Photon-Proton Collisions at HERA," *Z. Phys.*, vol. C70, pp. 17–30, 1996. 29, 30, 140

[37] C. Gwenlan, "Multijets in photoproduction at hera," *Acta Phys. Polon.*, vol. B33, pp. 3123–3128, 2002. 30, 31

[38] T. Sjostrand and P. Z. Skands, "Multiple interactions and the structure of beam remnants," *JHEP*, vol. 03, p. 053, 2004. 31, 32

[39] J. M. Butterworth and J. R. Forshaw, "Photoproduction of multi - jet events at HERA: A Monte Carlo simulation," *J. Phys.*, vol. G19, pp. 1657–1663, 1993. 36

[40] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, "Multiparton interactions in photoproduction at HERA," *Z. Phys.*, vol. C72, pp. 637–646, 1996. 36

[41] I. Abt *et al.*, "The H1 detector at HERA," *Nucl. Instrum. Meth.*, vol. A386, pp. 310–347, 1997. 42, 55, 56

[42] I. Abt *et al.*, "The Tracking, calorimeter and muon detectors of the H1 experiment at HERA," *Nucl. Instrum. Meth.*, vol. A386, pp. 348–396, 1997. 42, 52

[43] S. Eichenberger *et al.*, "A fast pipelined trigger for the H1 experiment based on multiwire proportional chamber signals," *Nucl. Instrum. Meth.*, vol. A323, pp. 532–536, 1992. 44

[44] H. Beck, "Principles and Operation of the z-Vertex Trigger," *H1 Internal Note, H1-05/96-479*, 1996. 44, 60

[45] J. Becker *et al.*, "A vertex trigger based on cylindrical multiwire proportional chambers," *Nucl. Instrum. Meth.*, vol. A586, pp. 190–203, 2008. 44

[46] D. Pitzl *et al.*, "The H1 silicon vertex detector," *Nucl. Instrum. Meth.*, vol. A454, pp. 334–349, 2000. 44

[47] B. List, "The H1 silicon tracker," *Nucl. Instrum. Meth.*, vol. A549, pp. 33–36, 2005. 44

[48] B. Andrieu *et al.*, "The H1 liquid argon calorimeter system," *Nucl. Instrum. Meth.*, vol. A336, pp. 460–498, 1993. 46

[49] R. D. Appuhn *et al.*, "The H1 lead/scintillating-fibre calorimeter," *Nucl. Instrum. Meth.*, vol. A386, pp. 397–408, 1997. 47

[50] H. Bethe and W. Heitler, "On the Stopping of fast particles and on the creation of positive electrons," *Proc. Roy. Soc. Lond.*, vol. A146, pp. 83–112, 1934. 48

[51] M. Jacquet, "The Steering of the H1 Level 2 Topological Trigger," *H1 Internal Note, H1-06/97-519*, 1997. 49

[52] A. Schoning, "A fast track trigger for the H1 collaboration," *Nucl. Instrum. Meth.*, vol. A518, pp. 542–543, 2004. 50

[53] H1 Collaboration, "H1REC: Reconstruction Program," *H1 Internal documentation*, 1996. 50

[54] The H1OO Group, "The H1 OO Physics Analysis Project," *H1 Internal documentation.* 53

[55] http://root.cern.ch/. 53

[56] B. Peez, M. Portheault and E. Sauvan, "An Energy Flow Algorithm for Hadronic Reconstruction In OO: Hadroo2," *H1 Internal Note, H1-01/05-616*, 2005. 53, 55

[57] Peez, M., *Recherche de déviations au Model Standard dans les processus de grande énergie transverse sur le collisionneur électron - proton HERA*. Ph.D. Dissertation, Université de Lyon, June 2003. 53

[58] Lee, West, "How to use the heavy flavour working group track, muon and electron selection code," *H1 Internal Sofware Manual*, 1997. 54

[59] Hellwig, Susanne, *Untersuchung der $D^*$ - $\pi_{slow}$ Double Tagging Methode in Charm-analysen.* Diplomarbeit, Universität Hamburg, Juni 2004. 55

[60] S. Osman, *Multiple Parton Interactions in Deep Inelastic ep-scattering at HERA*. Ph.D. Dissertation, Lund University, 2008. 57, 113

[61] T. Wolff *et al.*, "A Driftchamber Track Finder for the First Level Trigger of H1," *H1 Internal Note, H1-02/92-213*, 1992. 59

[62] J. Reidberger, "The H1 Trigger with Emphasis on Tracking Triggers," *H1 Internal Note, H1-01/95-419*, 1995. 59

[63] F. Sefkow, E. Elsen, H. Krehbiel, U. Straumann, and J. Coughlan, "Experience with the first level trigger of H1," *IEEE Trans. Nucl. Sci.*, vol. 42, pp. 900–904, 1995. 60

[64] S. Aid *et al.*, "Measurement of the Total Photon Proton Cross Section and its Decomposition at 200 GeV Centre of Mass Energy," *Z. Phys.*, vol. C69, pp. 27–38, 1995. 62

[65] G. D'Agostini, "A Multidimensional unfolding method based on Bayes' theorem," *Nucl. Instrum. Meth.*, vol. A362, pp. 487–498, 1995. 77, 91

[66] G. D'Agostini, "Comments on the Bayesian unfolding," *Zeus-Note-95-166*, 1995. 77, 91

[67] V. Blobel, "An unfolding method for high energy physics experiments," 2002. 77

[68] G. Cowan, "A survey of unfolding methods for particle physics," Prepared for Conference on Advanced Statistical Techniques in Particle Physics, Durham, England, 18-22 Mar 2002. 77

[69] F. D. Aaron *et al.*, "Charged Particle Production in High Q2 Deep-Inelastic Scattering at HERA," *Phys. Lett.*, vol. B654, pp. 148–159, 2007. 92

[70] A. Knutsson, A. Bacchetta, H. Jung, and K. Kutak, "An approach to fast fits of the unintegrated gluon density," 2008. 106

[71] Rick Field, "Talk presented at the fermilab me/mc tuning workshop, fermilab, october 4, 2002.," 113

[72] M. Y. Hussein, "A double parton scattering background to associate W H and Z H production at the LHC," *Nucl. Phys. Proc. Suppl.*, vol. 174, pp. 55–58, 2007. 113

# Acknowledgements

I would like to thank Prof.Dr.Beate Naroska and Dr.Carsten Niebuhr for the opportunity to come and work at DESY. With Prof.Dr.Naroska I also had the nice opportunity of having contact with the students at the Praktikum für Fortgeschrittene.

I want to thank very especially Hannes, my supervisor. His motivation for fundamental questions, help and commitment were crucial in many moments of this analysis.

My gratitude to my referees Prof.Dr.Elsen and Prof.Dr.Haller for agreeing to evaluate my work.

I am very thankful to Niladri and especially to Albert for reading, commenting and correcting my thesis.

Without the technical support (nearly 24h hours/day) and patience of Dr. Lucaci-Timoce this anlysis would not have been possible.

To the people from H1 and especially the HaQ group, the convenors, my referees, Dmitrij Ozerov and Arnd Specka, and the physics coordinators Andre Schoening and Cristinel Diaconu I would like to thank for their constructive comments. I am very pleased to share fruitful discussions on MPI with Sakar and Leif along these years.

Very important were also the people in the office: Axel, Nina, Zlatka and Tobias and the "Physics and cookies" team: Rico, Krzysztof, Michal, Gosta, Kaloyan, Svetlana and the "coffee" break team: Carlos, Sarah, Julia, Homer, Matt...

Ich danke ganz besonders Imke und Dominik, die immer bereit sind, mir zu helfen und Martin und Yvonne, die ich sehr hoch schätze. An Sibylle und die tolle Konzerte auch.

A la meva familia, Maria, Luis i Selma per estar sempre amb mi, escoltar-me i ajudar-me tant com han pogut.

和貴ちゃん、
いつも僕を支えてくれて、そして、
いつも新しい発見をありがとう。