

Determination of parton density functions using Monte Carlo event generators

Federico C. A. von Samson-Himmelstjerna
14. August 2009



Diplomarbeit

Fachbereich Physik
Freie Universität Berlin

Betreuer:
Prof. Dr. Paul Fumagalli

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 2 | The structure of the proton | 5 |
| 2.1 | Cross sections | 5 |
| 2.1.1 | Calculation of cross sections | 5 |
| 2.1.2 | Elastic scattering of unpolarized pointlike, spin- $\frac{1}{2}$ particles | 7 |
| 2.1.3 | Inelastic cross section | 9 |
| 2.2 | Quark parton model | 11 |
| 2.2.1 | Sum rules | 13 |
| 2.3 | QCD improved quark parton model | 14 |
| 2.3.1 | Factorized cross section | 14 |
| 2.3.2 | Sumrules | 21 |
| 2.3.3 | The DGLAP equations | 22 |
| 2.3.4 | DIS cross section | 23 |
| 3 | HERA, the H1-Detector and the measurement of F_2 | 25 |
| 3.1 | HERA | 25 |
| 3.2 | The H1 detector | 27 |
| 3.2.1 | Tracking system | 27 |
| 3.2.2 | Calorimeter | 29 |
| 3.2.3 | Muon detector | 32 |
| 3.3 | Measurement of F_2 | 32 |
| 4 | Monte Carlo Event Generators | 35 |
| 4.1 | Monte Carlo Event Generators | 36 |
| 4.2 | Parton showers | 37 |
| 4.3 | Hadronization | 38 |
| 4.4 | The Monte Carlo event generators of this work | 39 |
| 4.4.1 | PYTHIA | 39 |
| 4.4.2 | RAPGAP | 40 |

| | | |
|-----------|--|------------|
| 5 | Fitting | 41 |
| 5.1 | General formalism | 41 |
| 5.2 | Minimization | 42 |
| 5.3 | The SIMPLEX method | 43 |
| 5.3.1 | The SIMPLEX algorithm | 43 |
| 6 | Treatment of correlated systematic uncertainties | 47 |
| 6.1 | General ansatz | 49 |
| 6.2 | CTEQ-like treatment of correlated systematic uncertainties . . | 52 |
| 6.3 | Pascaud-like treatment of correlated systematic uncertainties | 53 |
| 7 | Data, software and Monte Carlo tuning | 55 |
| 7.1 | Dataset | 55 |
| 7.2 | Tools | 56 |
| 7.2.1 | HzTOOL | 56 |
| 7.2.2 | The routine <i>errortreat</i> | 56 |
| 7.3 | Histograms, bins and the determination of cross sections . . . | 58 |
| 7.4 | Tuning the Monte Carlo generators RAPGAP and PYTHIA | 59 |
| 8 | Determination of fitting parameters | 61 |
| 8.1 | The CTEQ6L PDF set | 61 |
| 8.2 | Selection of the fitting parameters | 62 |
| 9 | The influence of parton showers on PDFs | 69 |
| 9.1 | RAPGAP and PYTHIA excluding parton shower | 73 |
| 9.2 | RAPGAP including parton showers | 81 |
| 9.3 | PYTHIA including parton showers | 85 |
| 9.3.1 | k_{\perp} -ordered parton showers | 85 |
| 9.3.2 | Q^2 -ordered parton showers | 90 |
| 9.3.3 | Comparison with NLO PDF set | 94 |
| 9.3.4 | Comparison of χ^2 | 97 |
| 10 | Conclusions and Outlook | 99 |
| A | From the solution of the dirac equation to the two-body cross section | 105 |
| A.1 | Perturbative solution of the Dirac equation | 105 |
| A.2 | Electron propagator | 107 |
| A.3 | Photon propagator | 107 |

| | | |
|----------|--------------------------------------|------------|
| A.4 | Transition matrix | 108 |
| A.5 | Fermion-fermion scattering | 109 |
| A.6 | Two-body cross section | 110 |
| B | Fitted parameters | 113 |

Chapter 1

Introduction

In colliders the structure of matter is revealed by scattering experiments. The use of scattering experiments for the investigation of the structure of matter has its origin in the work of Rutherford et. al. at the beginning of the 20th century. In general, incoming objects with known momenta and energies are scattered on fixed or moving objects, called targets. By measuring the respective cross sections, information about the interaction between the object and the target as well as the structure of the target is obtained. Ideally, the incoming objects are pointlike with a well-known interaction, such as electrons or muons.

In the simplest case the target is pointlike as well. But this is often not the case. The deviation from a pointlike structure can then be described by so-called form factors. Form factors describe the spacial distribution of "charge" in the target and modify the measured cross sections.

This technique was used extensively by Hofstadter et. al. in the 50's and 60's to investigate nuclei. Here atomic form factors were introduced as a modification of the point structure. They are the Fourier transforms of the charge distribution of a nucleus.

When the investigation of protons began, it soon became clear that they are not pointlike. Thus, proton form factors were introduced to describe the deviation of the protons structure from a point.

The higher the energy transfer from the incoming object to the target, the higher is the resolution for measuring the target's structure. This fact is used in electron proton scattering, but it leads to a complication. Because of the high energy transfer, the proton breaks up - the elastic scattering becomes inelastic and inelastic structure functions are introduced to describe this.. From 1968 on, so-called scattering experiments at 20 GeV were made at the SLAC (**S**tanford **L**inear **A**cceleration **C**enter). At these energies of a magnitude before never reached, the incoming electron can probe deep into

the proton's structure, a process called Deep Inelastic Scattering (DIS). The results indicated pointlike structures within the proton. This was explained in Feynman's Quark Parton Model (QPM) (see sec. 2.2). The idea is that the proton consists of pointlike constituents - the so-called partons (which can be identified with the quarks). At sufficiently high energies these partons behave like free (Dirac) particles and are resolved. The electron then scatters at one of these constituents rather than at the entire proton. The inelastic structure function can then be expressed as the incoherent sum of elastic pointlike structure functions.

The distribution of the fraction of the proton's momentum over the partons is described by so-called Parton Distribution Functions (PDFs). They give the number of partons having a certain momentum fraction x of the proton. However the QPM is incomplete.

Quantum Chromodynamics (QCD) with the gluons had to be taken into account to expand the QPM into the QCD improved QPM.

The particle mediating the strong force, the gluon, is now taken into account. This has the effect, that partons can now radiate partons and by that change their momentum. Furthermore, scattering of gluons is considered now. With increasing energies more of these soft (low x) partons can be resolved and the PDFs become therefore energy-dependent, which leads to the observed scaling violation.

PDFs build the basis for most QCD calculations.

The world's only electron proton collider HERA opened from 1992 to 2007 the gate to test QCD at the previously not accessible kinematic regions of low momentum fractions x and high energy-momentum transfers Q^2 .

Important tools for theory and experiment, which need PDFs, are Monte Carlo event generators. They allow a computer simulation of events such as electron-proton collisions. All scattering products and the kinematics of these products are then accessible. The PDFs are an important input of a Monte Carlo event generator, since they describe the distribution of the partons momentum fraction x .

Parton radiation leads to the variety and the kinematics of the particles in the final state of an event.

In the QCD improved QPM the kinematics arising from parton radiation are usually not treated. This is alright when the focus is on calculating cross sections.

But, in a Monte Carlo event generator the complete final state is important. Therefore, it is crucial to keep track of the kinematics arising from parton radiation, since they determine the kinematics of the final state.

To treat the individual radiated particles, so-called parton shower are included in a Monte Carlo event generator. However, parton shower change the distribution of the momentum fraction x of the struck parton, which is given by the PDF.

In this work the influence of parton shower on the PDFs is investigated. To do so, PDFs are determined by fitting of Monte Carlo event generator-predictions for the structure functions F_2 to measurements from HERA. The resulting PDFs will be different from PDFs determined by other numerical methods not keeping track of parton kinematics.

The HERA data used in this work were taken in the years 1996/97 by the H1 collaboration and covers an x range of 3×10^{-4} to 2×10^{-2} and a Q^2 range from 1.5 GeV^2 to 150 GeV^2 . The data includes values for individual correlated systematic errors. The proper treatment of these errors plays an important role in the determination of the quality of the fit. In this work the routine *errortreat* was written to implement a proper treatment of the correlated systematic errors.

The influence of parton showers on the fitted PDFs was investigated for the Monte Carlo event generators PYTHIA and RAPGAP. The latter one is special, since it does not change the x distribution, although including parton shower. The fitted PDFs are compared with the PDF sets CTEQ6L and CTEQ6.1 M, which were determined by semi-analytical methods without parton shower.

The theoretical background, the tools and the results of this analysis are presented in this thesis.

In this work a system of natural units will be used, whereby $\hbar = c = 1$.

Chapter 2

The structure of the proton

Computing the (strong) color interactions of single gluons and quarks can be done using e.g. the Feynman rules. But experiments to probe these color interactions can, due to confinement, measure only interactions with entire hadrons instead of single gluons and quarks.

At HERA the structure of the proton was investigated in scattering experiments of electrons¹ and protons such as *H1* or *ZEUS*. Thereby a virtual photon γ^* with four momentum q is exchanged between the electron and the proton. One defines the quantity $Q^2 = -q^2$. With increasing virtuality Q^2 the photon resolves the structure of the proton in more detail which can be observed in the measured cross section. At sufficiently high Q^2 a transition from elastic scattering at the entire proton to an inelastic scattering at the constituents of the proton takes place. This observation was first described in the quark parton model and finally in its extension - the QCD improved quark parton model. In this chapter an overview of the determination of elastic and inelastic cross sections is given (sec. 2.1), as well as the main ideas of the quark parton model (sec. 2.2) and the QCD improved quark parton model (sec. 2.3).

2.1 Cross sections

2.1.1 Calculation of cross sections

The transition of an incoming wave function $\Phi_i(x)$ scattered at a 4-vector potential $A_\mu(x)$ into an outgoing wavefunction $\Phi_f(x)$ can be described by

¹The electron stands here for electrons and positrons. Both were used during the years.

the 1st-order transition matrix (see sec. A.4)

$$S_{if}^{(1)} = ie \int d^4x \bar{\Phi}_f(x') \gamma^\mu A_\mu(x') \Phi_i(x') \quad (2.1)$$

For a four body process $p_1 p_2 \rightarrow p_3 p_4$ this matrix can be written according to the Feynman rules (e.g. [8, 21, 13, 12]) as the product

$$S_{if}^{(1)} = \mathcal{M} \cdot (2\pi)^4 \delta^4(p_3 + p_4 - p_1 - p_2) \quad (2.2)$$

where \mathcal{M} is the so-called matrix element given by the propagator and the four vector currents of the particles².

The resulting cross section can be derived from Fermi's golden rule ([12, 20, 21]), which states, that the transition probability per unit time is

$$w = 2\pi \frac{d}{dt} |S_{if}|^2 \delta(p_i - p_f) \rho(E_f) \quad (2.3)$$

where $\rho(E) dE$ is the number states with Energy E in the energy interval $[E, E + dE]$.

Furthermore the cross section is defined as $\sigma = \frac{w}{j_{in}}$ and therefore $d\sigma = \frac{dw}{j_{in}}$. Thus the differential cross section depends on the product of the following quantities:

1. the square of the matrixelement (here the normalization $\frac{1}{\sqrt{V}}$ of each spinor is taken into account, which in any case cancels out later).

$$\frac{|\mathcal{M}|^2}{V^4}$$

2. the incoming current

$$|j_{in}| = \frac{|\vec{v}|}{V/2E_1}$$

3. the number of particles per unit volume in the target (because we have more than one target)

$$\frac{2E_2}{V}$$

4. the number of final states in the two body process

$$\frac{V}{2E_3} \frac{d^3p_3}{(2\pi)^3} \frac{V}{2E_4} \frac{d^3p_4}{(2\pi)^3}$$

²a brief derivation of the results eq. (2.1) and eq. (2.2) is given in the sections A.1 to A.5

5. the delta function (see eq. (2.2)) that guarantees energy-momentum conservation

$$(2\pi)^4 \delta^4(p_3 + p_4 - p_1 - p_2)$$

- 2) and 3) are combined to form the so-called *flux factor*³

$$\frac{\text{flux factor}}{V^2} = |\vec{v}| \frac{4E_1 E_2}{V^2} = \frac{4}{V^2} ((p_1 p_2)^2 - m_1^2 m_2^2)^{1/2} \quad (2.4)$$

- 4) and 5) are combined to form the Lorentz invariant phase factor

$$V^2 dL_{ips}(s, p_3, p_4) = V^2 (2\pi)^4 \delta^4(p_3 + p_4 - p_1 - p_2) \frac{d^3 p_3}{(2\pi)^3 2E_3} \frac{d^3 p_4}{(2\pi)^3 2E_4} \quad (2.5)$$

One can then write

$$d\sigma = |\mathcal{M}|^2 \frac{dL_{ips}}{\text{flux factor}} \quad (2.6)$$

2.1.2 Elastic scattering of unpolarized pointlike, spin- $\frac{1}{2}$ particles

Lets have a look now at the specific case of electron-muon scattering. Instead of the muon a dirac proton or any pointlike, spin- $\frac{1}{2}$ particle could be taken. The electron has the mass m_e and the muon (dirac proton) the mass $m \gg m_e$. The result will be important later considering electron parton scattering.

If the spin of the incoming and outgoing particles is not measured the spin-averaged square of \mathcal{M} has to be used instead, which is

$$\overline{|\mathcal{M}|^2} = \frac{1}{4} \sum_{s_1, s_2} \sum_{s_3, s_4} |\mathcal{M}|^2 = \left(\frac{e^2}{q^2}\right)^2 L_{\mu\nu} M^{\mu\nu} \quad (2.7)$$

³As the right side is relativistically invariant we can change to the laboratory frame, where this can easily be seen. Here:

$$\begin{aligned} p_1 &= (E_1, \vec{p}_1), p_2 = (m_2, 0) \Rightarrow p_1 p_2 = E_1 m_2 \\ &\Rightarrow ((p_1 p_2)^2 - m_1^2 m_2^2)^{1/2} = (E_1^2 - m_1^2)^{1/2} m_2 = |\vec{v}_1| E_1 E_2 \end{aligned}$$

with $\vec{v} = \vec{v}_1$ in the labframe

wherein the leptonic tensors are defined

$$L_{\mu\nu} := \frac{1}{2} \sum_{s_1 s_3} \bar{u}(p_3) \gamma_\mu u(p_1) \bar{u}(p_1) \gamma_\nu u(p_3) \quad (2.8)$$

$$M^{\mu\nu} := \frac{1}{2} \sum_{s_2 s_4} \bar{u}(p_4) \gamma^\mu u(p_2) \bar{u}(p_2) \gamma^\nu u(p_4) \quad (2.9)$$

"L" and "M" standing here for lepton and muon.

Using the trace theorems ([21],[13]) the leptonic tensors can be written

$$\begin{aligned} L_{\mu\nu} &= 2(p_{1\mu} p_{3\nu} + p_{1\nu} p_{3\mu} - (p_1 \cdot p_3) g_{\mu\nu}) \\ M^{\mu\nu} &= 2(p_2^\mu p_4^\nu + p_2^\nu p_4^\mu - (p_2 \cdot p_4) g^{\mu\nu}) \end{aligned}$$

In the fixed target frame the product of both tensors gives ([8])

$$L_{\mu\nu} M^{\mu\nu} = 16m^2 E_1 E_3 \left(\cos^2\left(\frac{\theta}{2}\right) + \frac{Q^2}{2m^2} \sin^2\left(\frac{\theta}{2}\right) \right)$$

where again $Q^2 = -q^2$. Eq. (2.6) is actually a multiple differential cross section, but is usually written $d\sigma$. Therefore one has to integrate over all quantities that are not observed. When looking for the differential cross section $\frac{d^2\sigma}{d\Omega dE_3}$ only the integration over p_4 is carried out. Thereby the following identity is used

$$\frac{d^3 \vec{p}_4}{2E_4} = \theta(E_4) \delta(p_4^2 - m^2) d^4 p_4$$

Integrating dL_{ips} now over p_4 gives overall 4-momentum conservation and leaves $\delta(p_4^2 - m^2)$. Since $p_4^2 = (q + p_2)^2 = m^2 + q^2 + 2p_2 q$ and $\nu = p_2 q$ one can rewrite the δ function:

$$\delta(p_4^2 - m^2) = \delta(q^2 + 2\nu) = \frac{1}{2} \delta\left(\nu - \frac{Q^2}{2}\right)$$

The differential $d^3 p_3$ still left in dL_{ips} can be rewritten as

$$\begin{aligned} d^3 p_3 &= |\vec{p}_3|^2 d|\vec{p}_3| d\Omega = |\vec{p}_3| E_3 dE_3 d\Omega \\ \Rightarrow \frac{d^3 p_3}{2E_3} &= \frac{1}{2} |\vec{p}_3| dE_3 d\Omega \approx \frac{1}{2} E_3 dE_3 d\Omega \end{aligned}$$

which results in

$$dL_{ips} = \frac{E_3 dE_3 d\Omega}{4(2\pi)^2} \delta\left(\nu - \frac{Q^2}{2}\right).$$

The flux factor is given by:

$$flux\ factor = 4\sqrt{(p_1 p_2)^2 - m_e^2 m^2} = 4\sqrt{m^2 E_1^2} = 4mE_1$$

in the last step $m^2 = p_2^2$, $m_e^2 \approx 0$ and, following from that, $p_1^2 = E_1^2 - m_e^2 \approx E_1^2$ was used.

Putting all this in (2.6) equates the differential cross section as:

$$\begin{aligned} \frac{d^2\sigma}{d\Omega dE_3} &= \frac{4\alpha^2 e_i^2}{Q^2} m E_3^2 \left(\cos^2\left(\frac{\theta}{2}\right) + \frac{Q^2}{2m^2} \sin^2\left(\frac{\theta}{2}\right) \right) \delta\left(\nu - \frac{Q^2}{2}\right) \\ &= \frac{\alpha^2 E_3}{q^4 E_1} 4m e_i^2 E_1 E_3 \left(\cos^2\left(\frac{\theta}{2}\right) + \frac{Q^2}{2m^2} \sin^2\left(\frac{\theta}{2}\right) \right) \delta\left(\nu - \frac{Q^2}{2}\right) \end{aligned} \quad (2.10)$$

2.1.3 Inelastic cross section

To investigate the proton structure in more detail the energy-momentum transfer to the proton, given by the virtuality Q^2 of the virtual photon, is increased. But there is a catch: due to the high energy transferred to the proton it will often break up and one can no longer assume a system with a proton coming in and a proton coming out. The bombarding lepton will break up the proton and only interact with a part of it (which will be identified by a so-called *parton*). This parton will form a jet and the rest of the proton, the so-called beam remnant, will form new hadrons (hadronize) which keep on flying more or less in the original beam direction. This process results in the proton losing its identity after the collision, while the lepton remains intact (see fig. 2.1).

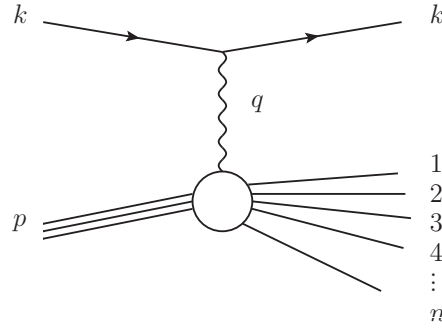


Figure 2.1: Inelastic e^- -p-scattering

For the calculation of the matrix element this means, that we do not know the details of the proton final state, but, at least, the leptonic part stays the same as in the elastic case. To tackle this problem in a system with the squared center of mass energy

$$s = (p + k)^2 \quad (2.11)$$

in addition to the four momentum of the virtual photon,

$$\begin{aligned} q &= k - k' \\ Q^2 &= -q^2, \end{aligned} \quad (2.12)$$

the following variables are defined:

$$x = \frac{Q^2}{2pq} \quad (2.13)$$

is called Bjorken x and will later, in the quark parton model (QPM), be identified with the fraction of the protons momentum taken by the interacting parton.

The inelasticity:

$$y = \frac{pq}{pk} \quad (2.14)$$

is a measure of the amount of energy transferred between the lepton and the hadron system. It is also related to the scattering angle in the center of mass frame.

An important relation resulting from these definitions is:

$$Q^2 = sxy \quad (2.15)$$

The center of mass energy of the γ^*p system is given by:

$$\hat{s} = (p + q)^2 \quad (2.16)$$

Next the matrix element is considered. In the elastic case the matrix element was expressed in form of a product of a leptonic tensor $L_{\mu\nu}^e$ and an elastic hadronic tensor $L_{\mu\nu}^p$ (see eq. (2.7)):

$$\mathcal{M} \propto L_{\mu\nu}^e L_{\mu\nu}^p$$

now to be replaced by an inelastic hadronic tensor $W_{\mu\nu}$

$$\mathcal{M} \propto L_{\mu\nu}^e W_{\mu\nu}$$

The hadronic tensor $W_{\mu\nu}$ is usually given in the parametrized form in terms of p , q and $g_{\mu\nu}$ ([8])

$$W_{\mu\nu} = W_1(-g_{\mu\nu} + \frac{q_\mu q_\nu}{q^2}) + W_2(p_\mu - \frac{p \cdot q}{q^2} q_\mu)(p_\nu - \frac{p \cdot q}{q^2} q_\nu)$$

The coefficients are so-called *structure functions* W_i . It can be shown ([8]) that the cross section then has the form:

$$\begin{aligned}
\frac{d^2\sigma}{d\Omega dE_3} &= \frac{\alpha^2}{4ME_1^2 \sin^4(\frac{\theta}{2})} (W_2(Q^2, \nu) \cos^2(\frac{\theta}{2}) + 2W_1(Q^2, \nu) \sin^2(\frac{\theta}{2})) \\
&= \frac{\alpha^2}{q^4} \frac{E_3}{E_1} \frac{4E_1 E_3}{M} (W_2(Q^2, \nu) \cos^2(\frac{\theta}{2}) + 2W_1(Q^2, \nu) \sin^2(\frac{\theta}{2}))
\end{aligned} \tag{2.17}$$

where M is the proton mass.

2.2 Quark parton model

Experimental measurements indicate that the proton consists of structureless particles, because for sufficiently high Q^2 the proton starts behaving like a free Dirac particle and the cross section (2.17) turns into (2.10). This inelastic ($\hat{s} \gg M^2$) scattering at high $Q^2 \gg M^2$ is called *deep inelastic scattering* (DIS). But this means that in this regime

$$\begin{aligned}
\frac{2W_1^{point}}{M} &= e_i^2 \frac{Q^2}{2m} \delta(\nu - \frac{Q^2}{2}) \\
\frac{W_2^{point}}{M} &= e_i^2 m \delta(\nu - \frac{Q^2}{2})
\end{aligned}$$

with m being the quark mass. Why this is so, will shortly become clear.

The fact that the scattering behaves at high Q^2 pointlike can be interpreted as follows: in this Q^2 -regime the substructure of the proton is resolved and the interaction time $\tau = \frac{1}{\nu}$ is short. So the virtual photon "sees" a single (high resolution), free (short interaction time) particle and it scatters elastically off it. These particles were called *partons* by Bjorken.

This is in contrast to elastic scattering where less energy is transferred and the interaction time is relatively long. The photon then "sees" only the blurred parton cloud comparable to the electron cloud of atoms.

The observation led to the *quark parton model* (QPM). In this model DIS is treated as the incoherent sum of point-like elastic scattering of spin- $\frac{1}{2}$ particles which are the constituents of the nucleon - the so-called *partons*.

The mathematical treatment of the parton model uses the approximation of the *infinite momentum frame* where $p \gg M$ (at HERA this frame is the *laboratory frame*). Here the partons can be seen as free, because at these high velocities the interaction time of the scattering is, due to time dilatation,

long compared to the time for interactions between the partons inside the proton. Assuming a parton carries the fraction x' of the protons momentum, then

$$\vec{p} = x' \vec{p}_{proton}$$

it then follows in the infinite momentum frame that

$$\begin{aligned} E_{parton} &= x' E_{proton} \\ \Rightarrow p^\mu &= x' p_{proton}^\mu. \end{aligned}$$

One can then define a *parton density function* (PDF) $f_j(x')$ such that $f_j(x')dx'$ gives the probability of finding a parton of type j with momentum fraction between x' and $x' + dx'$. The momentum distribution of a parton is given by the momentum weighted density $x' f_j(x')$. The kinematics in the *infinite momentum frame* are:

$$\begin{aligned} E_{parton} &= x' E_{proton} \\ &\Rightarrow \nu = x' \nu_{proton} \\ p_L &= x' p_{proton, L} \\ p_T &= p_{proton, T} = 0 \\ m^2 &= E_{parton}^2 - p^2 = x'^2 (E_{proton}^2 - p_{proton}^2) = x'^2 M^2 \end{aligned}$$

with longitudinal momentum p_L and transversal momentum p_T .

The structure functions for each parton expressed in terms of these variables are then

$$\frac{2W_1^{point}}{M} = e_i^2 \frac{Q^2}{2Mx'} \delta(x' \nu_{proton} - \frac{Q^2}{2}) \quad (2.18)$$

$$\frac{W_2^{point}}{M} = e_i^2 x' M \delta(x' \nu_{proton} - \frac{Q^2}{2}) \quad (2.19)$$

In this high Q^2 regime the W_i 's are usually replaced by their scaling limits F_1 , F_2 , which are defined such that they are dimensionless:

$$2F_1^{point} := 2W_1^{point} = e_i^2 \frac{Q^2}{2x'} \delta(x' \nu_{proton} - \frac{Q^2}{2}) = e_i^2 \frac{x}{x'} \delta(x' - x) \quad (2.20)$$

$$F_2^{point} := \frac{\nu_{proton} W_2^{point}}{M} = e_i^2 x' \delta(x' - x) \quad (2.21)$$

where x is Bjorken x .

As the partons can be assumed to be free one can simply add the contribu-

tions of the single partons incoherently and gets

$$2F_1(x) = \sum_i \int dx' e_i^2 f_i(x') \frac{x}{x'} \delta(x' - x) \quad (2.22)$$

$$F_2(x) = \sum_i \int dx' e_i^2 f_i(x') x' \delta(x' - x) \quad (2.23)$$

The delta distribution shows that one can identify Bjorken x with the momentum fraction x' . Thus, only a virtual photon with the right value for x can be absorbed by a parton that carries momentum fraction x' .

Accordingly, for high Q^2 one can write

$$F_1(x) = \frac{1}{2x} F_2(x) \quad (2.24)$$

$$F_2(x) = \sum_i e_i^2 x f_i(x) \quad (2.25)$$

The sum runs hereby only over the charged partons. Eq. (2.2) is the *Callan-Gross* relation. Eq. (2.25) predicts *Bjorken scaling* in the QPM, which is, that F_2 depends only on x and not on Q^2 .

2.2.1 Sum rules

There are constraints for the PDFs, the so-called *sum rules*, which are:

$$\int_0^1 dx u_v(x) \stackrel{!}{=} 2$$

$$\int_0^1 dx d_v(x) \stackrel{!}{=} 1$$

since $u_v(x)$ and $d_v(x)$ are the quark number distributions for the valence up-quark and the down-quark respectively and in the QPM one assumes 2 up-quarks and 1 down-quark in the proton.

Of course momentum conservation must be valid and, summing over all quarks, the fractions x' have to add up to 1

$$\sum_i \int dx' x' \cdot f_i(x') \stackrel{!}{=} 1. \quad (2.26)$$

But experimentally one finds

$$\sum_i \int_0^1 dx x \cdot f_i(x) \approx 0.5 \quad !$$

Furthermore: were the proton to consist of 3 quarks carrying all the momentum the PDFs should be δ -peaks⁴ at $x = \frac{1}{3}$.

The solution to these inconsistencies comes from QCD, that introduces gluons as the carriers of the strong force. They carry the rest of the momentum and can split into $q\bar{q}$ -pairs - the so-called sea quarks (in contrast to the 3 valence quarks, 2 u and 1 d in the proton). The next section treats the QCD improved QPM.

2.3 QCD improved quark parton model

The basis of any QCD calculations lies in the factorization theorem which says that the cross section can be factorized in a parton density function and a hard scattering cross section. In the following the hard scattering cross section will be calculated.

2.3.1 Factorized cross section

Using the factorization theorem the γ^*p cross section takes the form:

$$\begin{aligned} \left(\sigma(x, Q^2)\right)_{\gamma^*p} &= \sum_i \int_0^1 dz \int_0^1 d\xi f_i(\xi) \delta(x - z\xi) \left(\hat{\sigma}(z, Q^2)\right)_{\gamma^*i} \\ &= \sum_i \int_0^1 \frac{d\xi}{\xi} f_i(\xi) \left(\hat{\sigma}\left(\frac{x}{\xi}, Q^2\right)\right)_{\gamma^*i} \end{aligned} \quad (2.27)$$

Note: the left side is the γ^* -proton cross section (subscript γ^*p), while on the right side $\hat{\sigma}$ is the γ^* -parton cross section (subscript γ^*i).

One can understand this cross section in the following example of a single gluon emission ($\gamma^*q \rightarrow gq$).

In the γ^* -parton interaction a parton with initial momentum $p_i = \xi p$ interacts with a virtual photon of 4-momentum q . The parton can radiate a gluon and pass some momentum, e.g. a fraction $(1 - z)$ of its initial momentum p_i , to the radiated gluon. It then carries on with a changed

⁴Due to motion of the quarks inside the proton, the so-called Fermi motion, it would rather be a Gauss-peak centered at $x = \frac{1}{3}$

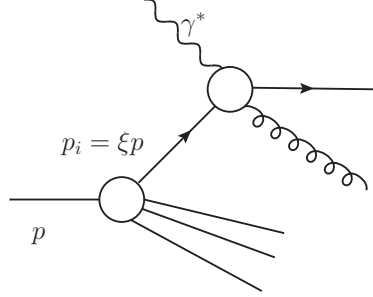


Figure 2.2: Single gluon emission

momentum $p'_i = zp_i = z\xi p_{prot}$, while the radiated gluon has a momentum $(1-z)p_i = (1-z)\xi p$.

$(\hat{\sigma}(z, Q^2))_{\gamma^*i}$ is the cross section for the absorption of a photon of momentum q by a parton of momentum fraction p_i . This cross section depends, completely analogous to the γ^* -proton case, on the quantity $z = \frac{Q^2}{2p_i q}$ (γ^* -proton: $x = \frac{Q^2}{2p_{prot} q}$). In the γ^* -proton interaction it was a virtual photon of momentum q that interacted with a parton carrying the appropriate momentum $p = xp_{prot}$. Here, in the γ^* -parton interaction, a virtual photon of momentum q interacts with a parton carrying the appropriate momentum $p'_i = zp_i (= x\xi p_{prot})$.

The two processes are related to each other by eq. (2.27).

In eq. (2.27) the cross section $(\sigma(x, Q^2))_{\gamma^*p}$ is calculated at a specific x by integrating over all those processes resulting in a momentum fraction $z\xi = x$. The latter condition is reflected by the delta function and an intergration over all ξ and z in eq. (2.27). This takes into account all combinations of ξ (momentum fraction before the radiation) and z (momentum fraction carried away by gluon radiation) that lead to a resulting momentum fraction x (this is assured by the delta function) and weights them by the probability $f_i(\xi)d\xi$ of finding a parton with initial momentum fraction ξ .

The structure functions are related to the cross sections for virtual photons of transverse and longitudinal polarization ([8]) by

$$\begin{aligned}\sigma_T &= \frac{4\pi\alpha}{s} 2F_1 =: \sigma_0 2F_1 \\ \sigma_L &= \frac{4\pi\alpha}{s} \left[\frac{F_2}{x} - 2F_1 \right] =: \sigma_0 \left[\frac{F_2}{x} - 2F_1 \right]\end{aligned}\tag{2.28}$$

where $\sigma_0 = \frac{4\pi^2\alpha}{s}$.

F_2 is then calculated via

$$\frac{F_2}{x} = \frac{\sigma_T + \sigma_L}{\sigma_0}\tag{2.29}$$

It can be shown ([8]), that at γ^* -parton level the cross sections for a single massless quark of charge e_i interacting with a γ^* have the form

$$\begin{aligned}\hat{\sigma}_T &= e_i^2 \hat{\sigma}_0 \delta(1-z) \\ \hat{\sigma}_L &= 0\end{aligned}\quad (2.30)$$

Substituting this in (2.27), summing over all quarks and using (2.29) gives:

$$\begin{aligned}\frac{F_2(x)}{x} &= \sum_i \int_x^1 \frac{d\xi}{\xi} q_i(\xi) e_i^2 \delta\left(1 - \frac{x}{\xi}\right) \\ &= \sum_i e_i^2 q_i(x)\end{aligned}\quad (2.31)$$

This reproduces the result of the QPM model (2.25) and is the zeroth-order contribution to the structure function.

Turning to the $\mathcal{O}(\alpha_s)$ contribution to the hard processes:

One process that has to be added to the QPM is gluon radiation by quarks (QCD Compton process) and is shown in fig. 2.3.

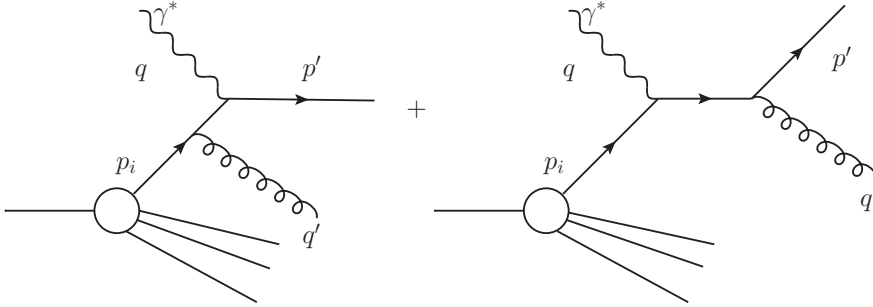


Figure 2.3: Gluon radiation by quarks

The following formulae are outlined and derived in detail in [13] or [8]. The QCD compton process $\gamma^*q \rightarrow qq$ can be calculated in analogy to the QED compton process $\gamma^*e \rightarrow \gamma e$.

This gives for the matrix element squared([13]):

$$\overline{|\mathcal{M}|^2} = 32\pi^2 e_i^2 \alpha \alpha_s \frac{4}{3} \left(-\frac{\hat{t}}{\hat{s}} - \frac{\hat{s}}{\hat{t}} + \frac{2\hat{u}Q^2}{\hat{s}\hat{t}} \right) \quad (2.32)$$

where

$$\begin{aligned}\hat{s} &= (q + p_i)^2 = Q^2 \frac{1-z}{z} \quad \text{since } z = \frac{Q^2}{2qp_i} = \frac{Q^2}{\hat{s}^2 + Q^2} \\ \hat{u} &= (q - q')^2 \\ \hat{t} &= (q - p')^2\end{aligned}$$

For

$$\hat{t} \ll \hat{s} \quad (2.33)$$

it can be approximated by

$$|\overline{\mathcal{M}}|^2 = 32\pi^2 e_i^2 \alpha \alpha_s \frac{4}{3} \left(-\frac{\hat{s}}{\hat{t}} + \frac{2\hat{u}Q^2}{\hat{s}\hat{t}} \right)$$

Using

$$\begin{aligned} s &= Q^2 \frac{(1-z)}{z} \\ s+t+u &= -Q^2 \end{aligned}$$

gives

$$|\overline{\mathcal{M}}|^2 = 32\pi^2 e_i^2 \alpha \alpha_s \left(-\frac{1}{\hat{t}} \right) Q^2 \frac{4}{3} \frac{1+z^2}{z(1-z)}$$

Together with the formula for the virtual photon differential cross section eq. (A.30) (replacing $s^2 \rightarrow \hat{s}(\hat{s} + q^2)$ in case of a virtual photon)

$$\frac{d\hat{\sigma}}{dt} = \frac{1}{16\pi\hat{s}(\hat{s} + Q^2)} |\overline{\mathcal{M}}|^2 \quad (2.34)$$

and $\frac{d\sigma}{dk_\perp^2} = \frac{d\sigma}{dt} \frac{dt}{dk_\perp^2} = -(1-z) \frac{d\sigma}{dt}$ this yields

$$\frac{d\hat{\sigma}}{dk_\perp^2} \simeq \sum_i e_i^2 \hat{\sigma}_0 \frac{1}{k_\perp^2} \frac{\alpha_s}{2\pi} P_{qq}(z) \quad (2.35)$$

where

$$P_{qq}(z) = \frac{4}{3} \frac{1+z^2}{1-z}$$

$P_{qq}(z)$ represents the probability of a quark emitting a gluon of momentum fraction $(1-z)$ in becoming a quark of momentum fraction z . This differential cross section has two divergencies. One for $z \rightarrow 1$, which corresponds to emission of a soft massless gluon. It is therefore called infrared divergency and is cancelled by virtual corrections.

The other divergency occurs for $k_\perp \rightarrow 0$. As k_\perp is related to the scattering angle this divergency is called collinear. It will be absorbed in the PDFs.

The total cross section is now obtained by integrating (2.35)

$$\begin{aligned}
 \hat{\sigma}(z, Q^2) &= \int_0^{k_{\perp}^2(max)} \frac{d\hat{\sigma}}{dk_{\perp}^2} dk_{\perp}^2 \\
 &\simeq \lim_{\kappa \rightarrow 0} \sum_i e_i^2 \hat{\sigma}_0 \frac{\alpha_s}{2\pi} P_{qq}(z) \int_{\kappa}^{k_{\perp}^2(max)} \frac{dk_{\perp}^2}{k_{\perp}^2} + \dots \\
 &\simeq \lim_{\kappa \rightarrow 0} \sum_i e_i^2 \hat{\sigma}_0 \frac{\alpha_s}{2\pi} P_{qq}(z) \left(\ln\left(\frac{Q^2}{\kappa^2}\right) + \ln\left(\frac{1-z}{4z}\right) \right) + \dots
 \end{aligned} \tag{2.36}$$

To regulate the divergency for $k_{\perp}^2 \rightarrow 0$ the cut-off κ is introduced at the lower limit. Furthermore it can be shown, that $k_{\perp, max}^2 = \frac{\hat{s}}{4} = \frac{Q^2(1-z)}{4z}$ ([8, 13]). This can now be added to the zeroth order QPM parton cross section (2.30) to calculate the structure functions, which gives:

$$\begin{aligned}
 \frac{F_2(x, Q^2)}{x} &= \\
 &\sum_i e_i^2 \int_x^1 \frac{d\xi}{\xi} q_i(\xi) \left(\delta\left(1 - \frac{x}{\xi}\right) + \frac{\alpha_s}{2\pi} P_{qq}\left(\frac{x}{\xi}\right) \ln\left(\frac{Q^2}{\kappa^2}\right) + \frac{\alpha_s}{2\pi} C\left(\frac{x}{\xi}\right) \right)
 \end{aligned} \tag{2.37}$$

where $C(z) = \ln\left(\frac{1-z}{4z}\right) + \dots$

The structure function depends now on x and Q^2 and not just on x alone, as it was the case in the parton model! Thus, the scaling prediction of the parton model is violated due to taking gluon emission into account.

The divergent part $\kappa \rightarrow 0$ is absorbed into a modified quark probability distribution. Therefore the parton densities are redefined.

The idea is to separate the perturbative part from the non-perturbative, singular part by introducing the arbitrary collinear factorization scale $\mu^2 \gg \kappa^2$. All $k_{\perp}^2 < \mu^2$ with the singularity ($k_{\perp} \rightarrow 0$) are now absorbed into a new renormalized parton density $q_i(x, \mu^2)$. The non-singular part with $k_{\perp} > \mu$ can be determined perturbatively.

Mathematically one simply uses

$$\ln\left(\frac{Q^2}{\kappa^2}\right) = \ln\left(\frac{Q^2}{\mu^2}\right) + \ln\left(\frac{\mu^2}{\kappa^2}\right)$$

and then defines

$$q_i(x, \mu^2) = q_i^0(x) + \frac{\alpha_s}{2\pi} \int_x^1 \frac{d\xi}{\xi} q_i^0(\xi) P_{qq}\left(\frac{x}{\xi}\right) \ln\left(\frac{\mu^2}{\kappa^2}\right) \tag{2.38}$$

resulting in

$$\frac{F_2^q(x, Q^2)}{x} = \sum_i e_i^2 \int_x^1 \frac{d\xi}{\xi} q_i(\xi, \mu^2) \left[\delta\left(1 - \frac{x}{\xi}\right) + \frac{\alpha_s}{2\pi} P_{qq}\left(\frac{x}{\xi}\right) \ln\left(\frac{Q^2}{\mu^2}\right) + \frac{\alpha_s}{2\pi} C\left(\frac{x}{\xi}\right) \right] \quad (2.39)$$

The last term is only correct to order α_s . An exact definition of the parton density and the non-singular piece depends on the renormalization and the factorization scheme⁵.

The renormalized PDF cannot be calculated perturbatively, but its variation with $\ln(\mu^2)$ can be derived from (2.38)

$$\frac{\partial q_i(x, \mu^2)}{\partial \ln(\mu^2)} = \frac{\alpha_s}{2\pi} \int_x^1 \frac{d\xi}{\xi} q_i(x, \mu^2) P_{qq}\left(\frac{x}{\xi}\right) \quad (2.40)$$

The second additional hard process beyond the QPM is the $\mathcal{O}(\alpha_s)$ process of quark-antiquark production by gluons or boson-gluon fusion (BGF) shown in fig. 2.4.

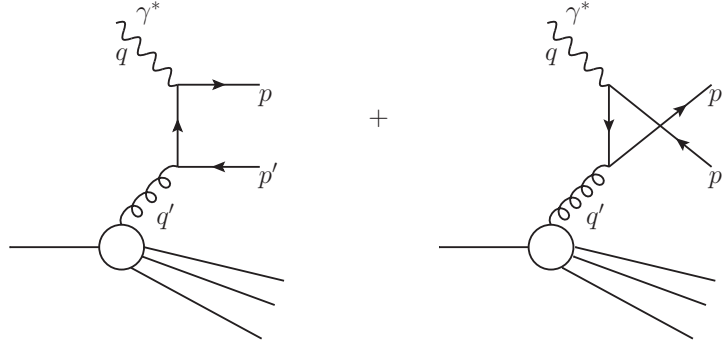


Figure 2.4: Boson gluon fusion

The mean square of the matrix element can be obtained from (2.32) by crossing and is (see [13])

$$|\overline{\mathcal{M}}|^2 = 32\pi^2 e_i^2 \alpha_s \frac{1}{2} \left(-\frac{\hat{u}}{\hat{t}} - \frac{\hat{t}}{\hat{u}} + \frac{2\hat{s}Q^2}{\hat{t}\hat{u}} \right) \quad (2.41)$$

⁵In the so-called DIS scheme, for example, the $C(z)$ term is absorbed into the renormalized PDF. With the choice $\mu^2 = Q^2$ as factorization scale (2.39) becomes

$$F_2^q(x, Q^2) = \sum_i e_i^2 x q_i(x, Q^2)$$

which looks like the QPM result just with $q_i(x) \rightarrow q_i(x, Q^2)$.

which gives, once more together with (2.34) and after integrating over k_{\perp}^2 (compare derivative of (2.36)),

$$\hat{\sigma} \simeq \sum_i e_i^2 \hat{\sigma}_0 \frac{\alpha_s}{2\pi} P_{qg}(z) \left(\ln\left(\frac{Q^2}{\kappa^2}\right) + \ln\left(\frac{1-z}{4z}\right) \right). \quad (2.42)$$

As before, κ was introduced to regulate the divergency for $k_{\perp} \rightarrow 0$ and

$$P_{qg} := \frac{1}{2} [z^2 + (1-z)^2]$$

is the splitting function that gives the probability for the process $g \rightarrow q(z)\bar{q}(1-z)$.

This process contributes to the structure function F_2

$$\frac{F_2^g(x, Q^2)}{x} = \sum_i e_i^2 \int_x^1 \frac{d\xi}{\xi} g^0(\xi) \frac{\alpha_s}{2\pi} P_{qg}\left(\frac{x}{\xi}\right) \ln\left(\frac{Q^2}{\kappa^2}\right) \quad (2.43)$$

and has to be added to (2.39).

And, as before, the singular part is separated by the factorization scale μ^2 and by renormalizing $g^0(x)$ to $g(x, \mu^2)$.

Including BGF in (2.40) gives then the final expression:

$$\boxed{\frac{\partial q_i(x, \mu^2)}{\partial \ln \mu^2} = \frac{\alpha_s}{2\pi} \int_x^1 \frac{d\xi}{\xi} \left[q_i(x, \mu^2) P_{qq}\left(\frac{x}{\xi}\right) + g(x, \mu^2) P_{qg}\left(\frac{x}{\xi}\right) \right].} \quad (2.44)$$

This type of equation, describing the variation of a parton density, is called an *evolution equation*. In the next section a special form of evolution equations, the DGLAP evolution equations ((2.44) belongs to them), will be introduced.

The essential idea of improving the QPM is adding processes via splitting functions such as $P_{qq}(z)$ and $P_{gq}(z)$. The procedure of how to achieve this is given in the derivation of (2.44). To have a complete set of equations one needs to take two more processes into account in therefore getting two more splitting functions $P_{gq}(z)$ and $P_{gg}(z)$ (the 4 splittings are summarized in fig. 2.5). They stand for the processes $q \rightarrow g(z)q(1-z)$ and $g \rightarrow g(z)g(1-z)$. The first one, $P_{gq}(z)$ can simply be derived by the fact that $P_{gq}(z) \stackrel{!}{=} P_{qq}(1-z)$, which yields:

$$P_{gq}(z) = \frac{4}{3} \frac{1 + (1-z)^2}{z} \quad (2.45)$$

Deriving P_{gg} is a bit more tricky and the result is simply stated

$$P_{gg}(z) = 6 \left[\frac{1-z}{z} + \frac{z}{1-z} + z(1-z) \right] \quad (2.46)$$

The Feynman graphs of the four splitting functions are shown in fig. 2.5

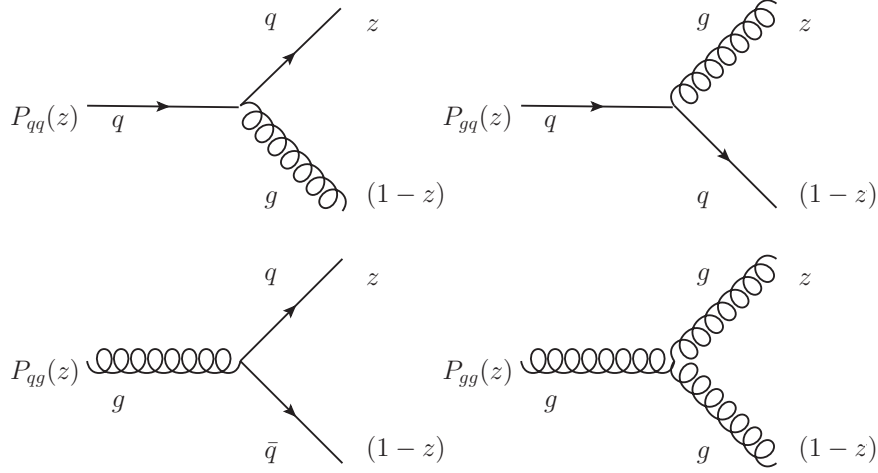


Figure 2.5: The splitting functions

2.3.2 Sumrules

The QCD Lagrangian conserves fermion number and flavour. To fulfill these constraints the PDFs have to obey again sumrules for the renormalized PDFs - at least to the order of the foregoing calculations which is $\mathcal{O}(\alpha_s)$. These rules are:

$$\int_0^1 dx [q_i(x, Q^2) - \bar{q}_i(x, Q^2)] = v_i \quad (2.47)$$

where $v_i = 2, 1, 0, \dots$ for the u, d, s, ... flavours in the proton.

As the overall momentum will be conserved the following must be valid:

$$\int_0^1 dx x \left[\sum_i (q_i(x, Q^2) + \bar{q}_i(x, Q^2)) + g(x, Q^2) \right] = 1 \quad (2.48)$$

i runs here over all active flavours and it is the the proper extension of (2.26). Conservation of momentum and flavour gives some constraints for the splitting functions as well:

$$\int_0^1 dz z [P_{qq}(z) + P_{gq}(z)] = 0 \quad (2.49)$$

$$\int_0^1 dz z [2n_f P_{qg}(z) + P_{gg}(z)] = 0 \quad (2.50)$$

$P_{qq}(z)$ has a singularity for $z \rightarrow 1$. Therefore loop corrections have to be taken into account to correct for this singularity. To do this the following ansatz for a corrected $P_{qq}(z)$ is made:

$$P_{qq}(z) = P_{qq}^+(z) + k\delta(1-z) \quad (2.51)$$

Eq. (2.38) can be written as

$$\begin{aligned} q_i(x, \mu^2) &= \int_x^1 \frac{d\xi}{\xi} q_i^0\left(\frac{x}{\xi}\right) \left(\delta\left(1 - \frac{x}{\xi}\right) + \frac{\alpha_s}{2\pi} P_{qq}\left(\frac{x}{\xi}\right) \ln\left(\frac{\mu^2}{\kappa^2}\right) \right) \\ &= \int_x^1 \frac{d\xi}{\xi} q_i^0\left(\frac{x}{\xi}\right) \left(\delta\left(1 - \frac{x}{\xi}\right) + \frac{\alpha_s}{2\pi} \left(P_{qq}^+\left(\frac{x}{\xi}\right) + k\delta\left(1 - \frac{x}{\xi}\right) \right) \ln\left(\frac{\mu^2}{\kappa^2}\right) \right) \end{aligned} \quad (2.52)$$

$\left(\delta\left(1 - \frac{x}{\xi}\right) + \frac{\alpha_s}{2\pi} P_{qq}\left(\frac{x}{\xi}\right) \ln\left(\frac{\mu^2}{\kappa^2}\right)\right)$ is the probability that a quark goes into a quark. Quark number conservation demands therefore

$$\int_0^1 dz \left(\delta(1 - z) + \frac{\alpha_s}{2\pi} (P_{qq}^+(z) + k\delta(1 - z)) \ln\left(\frac{\mu^2}{\kappa^2}\right) \right) = 1 \quad (2.53)$$

which can only be true when

$$\frac{\alpha_s}{2\pi} \int_0^1 dz (P_{qq}^+(z) + k\delta(1 - z)) = 0 \quad (2.54)$$

To tackle the singularity the "+"-distribution is introduced:

$$\int_0^1 dx \frac{f(x)}{(1-x)_+} = \int_0^1 dx \frac{f(x) - f(1)}{1-x} \quad (2.55)$$

Using this for $P_{qq}^+(z)$ in 2.54 gives then $k = \frac{3}{2}$.

This leads to the final expressions for $P_{qq}(z)$ (2.51) and $P_{gg}(z)$ (not derived):

$$P_{qq}(z) = \frac{4}{3} \left[\frac{1+z^2}{(1-z)_+} + \frac{3}{2} \delta(1-z) \right] \quad (2.56)$$

$$P_{gg}(z) = 6 \left[\frac{1-z}{z} + \frac{z}{(1-z)_+} + z(1-z) \right] + \frac{33-2n_f}{6} \delta(1-z) \quad (2.57)$$

2.3.3 The DGLAP equations

The formalism described in the previous section was developed over a number of years by Dokshitzer, Gribov, Lipatov, Altarelli and Parisi and led to four coupled equations known as the DGLAP equations (an acronym for the five names) ([2, 9, 10, 11, 18]).

They can be summarized in a matrix form

$$\begin{aligned} \frac{\partial}{\partial \ln Q^2} \begin{pmatrix} q_i(x, Q^2) \\ g(x, Q^2) \end{pmatrix} &= \frac{\alpha_s(Q^2)}{2\pi} \sum_k \int_x^1 \frac{d\xi}{\xi} \\ &\begin{pmatrix} P_{q_i q_k}\left(\frac{x}{\xi}, \alpha_s(Q^2)\right) & P_{q_i g}\left(\frac{x}{\xi}, \alpha_s(Q^2)\right) \\ P_{g q_k}\left(\frac{x}{\xi}, \alpha_s(Q^2)\right) & P_{g g}\left(\frac{x}{\xi}, \alpha_s(Q^2)\right) \end{pmatrix} \begin{pmatrix} q_k(\xi, Q^2) \\ g(\xi, Q^2) \end{pmatrix} \end{aligned} \quad (2.58)$$

where the $q_i(x, Q^2)$, $q_k(x, Q^2)$ can be quark or antiquark distributions. The splitting functions are expanded as power series in α_s :

$$\begin{aligned} P_{q_i q_k}(z, \alpha_s) &= \delta_{ik} P_{qq}^{(0)} + \frac{\alpha_s}{2\pi} P_{q_i q_k}^{(1)(z)} + \dots \\ P_{qg}(z, \alpha_s) &= P_{qg}^{(0)} + \frac{\alpha_s}{2\pi} P_{qg}^{(1)(z)} + \dots \\ P_{gq}(z, \alpha_s) &= P_{gq}^{(0)} + \frac{\alpha_s}{2\pi} P_{gq}^{(1)(z)} + \dots \\ P_{gg}(z, \alpha_s) &= P_{gg}^{(0)} + \frac{\alpha_s}{2\pi} P_{gg}^{(1)(z)} + \dots \end{aligned}$$

The δ_{ik} in front of the leading order term of P_{qq} stems from the fact that in a single quark gluon vertex flavour is conserved in leading order. The splitting functions of the foregoing section can be identified with the leading order terms $P^{(0)}$ and the sumrules they obey refer to these 0th-order $P^{(0)}$.

The DGLAP equations give a formalism to calculate the evolution of the PDFs with $\ln(Q^2)$ from a starting scale Q_0^2 but they do not allow the calculation of the PDFs at the starting scale Q_0^2 . The PDF at Q_0^2 must be given by non-perturbative methods in form of a parametrization of the x dependence of the PDF at Q_0^2 .

2.3.4 DIS cross section

Using the renormalized PDFs, the differential cross section (2.17) can now be expressed in x, y, Q^2 ([8]):

$$\begin{aligned} \frac{d^2\sigma}{dx dQ^2} &= \frac{2\pi\alpha^2}{xQ^4} ((1 + (1 - y)^2)(F_2(x, Q^2) - y^2 F_L(x, Q^2)) \\ &= \frac{d^2\sigma}{dx dQ^2} = \frac{2\pi\alpha^2 Y_+}{xQ^4} \cdot (F_2(x, Q^2) - \frac{y^2}{Y_+} F_L(x, Q^2)) \end{aligned} \quad (2.59)$$

with the longitudinal structure function $F_L = F_2 - 2xF_1$ and $Y_+ = 1 + (1 - y)^2$. Eq. (2.59) is the key equation for measurements done in this thesis.

Chapter 3

HERA, the H1-Detector and the measurement of F_2

In this chapter the ep collider HERA at the DESY laboratory in Hamburg, the H1 detector and the measurement of the structure function F_2 are introduced.

3.1 HERA

The HERA (**H**adron **E**lektron **R**ing **A**nlage), shown in fig. 3.1, is the first electron-proton collider. It consists of two independent rings of a circumference of 6.3 km - one ring accelerating and storing electrons in one direction and one accelerating and storing protons in the other.

Before the particles are injected into the HERA, they are pre-accelerated at different stations.

The acceleration of the proton starts in the H^- -LINAC where negatively charged hydrogen ions are brought to 50 MeV. The electrons are stripped off the ions and the latter ones are then injected into the DESY III ring in form of 11 bunches. There they are accelerated up to 7.5 GeV with a crossing distance equal to HERA's (which is 96 ns). Then the bunches are injected into the PETRA (**P**ositron **E**lektron **T**andem **R**ing**A**nlage) where they are brought to 40 GeV and then injected into HERA. The procedure is repeated until HERA is filled with 210 bunches. HERA then accelerates the protons to their final energy of 920 GeV.

The electron or positrons are first accelerated with the LINAC I or II respectively to 450 MeV. Bunch by bunch the leptons are injected into the

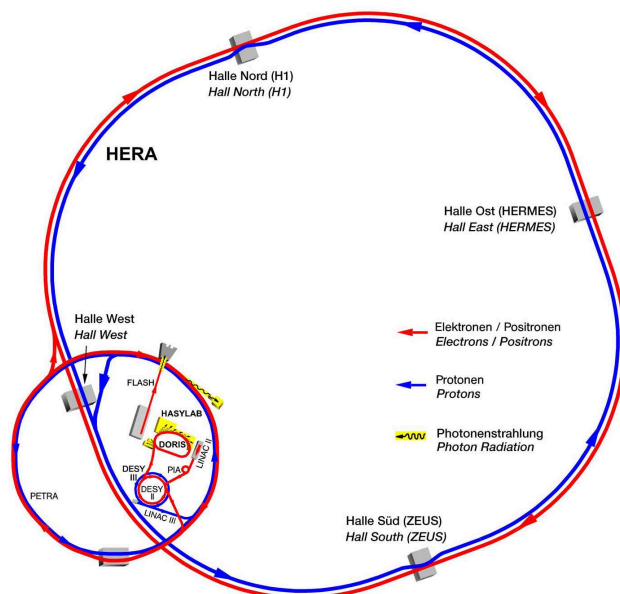


Figure 3.1: The HERA collider

DESY II ring and accelerated to 7 GeV until a number of 70 bunches is reached. These are then injected with a distance of 96 ns into PETRA II and there accelerated to 14 GeV. Finally the 210 bunches are injected into HERA and accelerated to 27.6 GeV. Every 96 ns the bunches cross each other (this corresponds to a distance of 28.8 m or a crossing frequency of ≈ 10 MHz) but not every bunch has a collision partner. Bunches without a partner are called pilot bunches and are used to determine the beam induced detector background.

Every bunch consists either of 0.4×10^{11} electrons or of 0.7×10^{12} protons.

The protons are kept on track by superconducting dipole magnets reaching 4.68 T, while the much lighter electrons only need normal conducting magnets reaching a field strength of 0.164 T.

There are four experiments installed at HERA. The two collider experiments H1 and ZEUS and the two fixed target experiments HERMES and HERA-B. In the following the H1-Detector will be described, on which the data for this work were recorded.

3.2 The H1 detector

The H1 detector is a multi-purpose detector that measures direction, momentum, energy and charge of particles resulting from ep collisions. It provides almost complete coverage of the interaction region. Only the part covered by the beam pipe is not accessible for the measurement.

The main detector components are shown in fig. 3.2. In all the detector has a size of $12m \times 10m \times 15m$ and a weight of approx. 2800 tons. The protons as shown in fig. 3.2 come in from the right with 920 GeV^1 and the electrons from the left with 27.6 GeV . All in all the center of mass energy is $\sqrt{s} = 319 \text{ GeV}$. Two main regions of the detector can be distinguished - the tracking system and the calorimeters.

The tracking system measures the tracks, the momenta and the charges of the collision products. The tracks of charged particles are bent by a homogeneous magnetic field of 1.15 T provided by an superconducting magnet. The curvature of the charged particles in the magnetic field allows their charge and momentum to be determined.

In the calorimeters located in the same magnetic field the energies are measured (in principle tracks can and will be measured here as well).

3.2.1 Tracking system

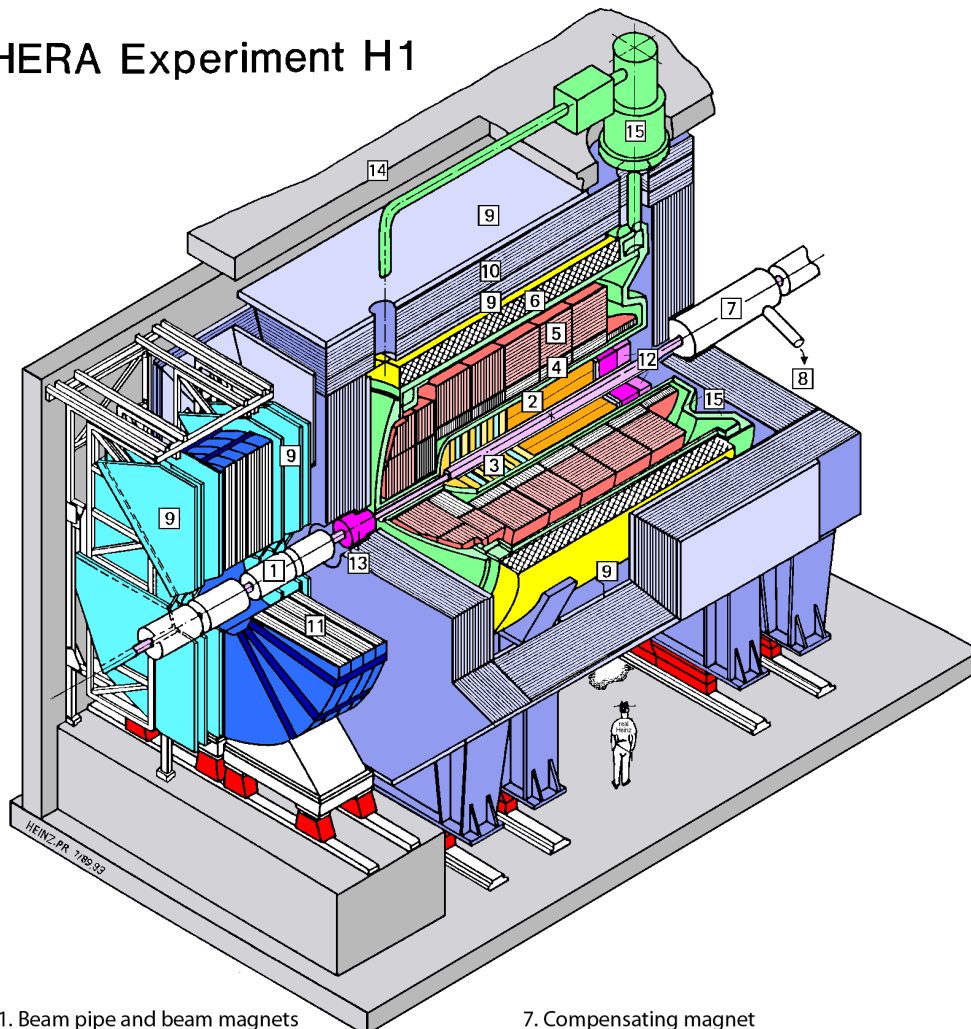
The tracking system (fig. 3.3 and fig. 3.4) is located in the inner part of the detector consisting of two parts - the central tracking detector (CTD) and the forward tracking detector (FTD). The former consists of the central silicon tracker (CST), two central jet chambers (CJC 1 and CJC 2), the central inner z-chamber (CIZ) and the central outer z-chamber (COZ).

Central silicon tracker

The innermost part of the detector is formed by the central silicon tracker. It consists of two concentric layers of silicon sensors, that measure tracks of charged particles close to the interaction vertex with high accuracy. This allows the decay length of charmed or beauty hadrons, which is in the range of a few hundred micrometers, to be measured.

¹before 1998 it was 820 GeV

HERA Experiment H1



- | | |
|--|---------------------------|
| 1. Beam pipe and beam magnets | 7. Compensating magnet |
| 2. Central tracking chambers | 8. Helium cryogenics |
| 3. Forward tracking and Transiton radiators | 9. Muon chambers |
| 4. Electromagnetic calorimeter (lead, liquid argon) | 10. Return yoke |
| 5. Hadronic calorimeter (stainless steel) | 11. Muon-Toroid-Magnet |
| 6. Superconducting coil (1.15 T) | 12. SpaCal |
| | 13. Plug calorimeter |
| | 14. Concrete screen |
| | 15. Liquid Argon cryostat |

Figure 3.2: The H1 detector

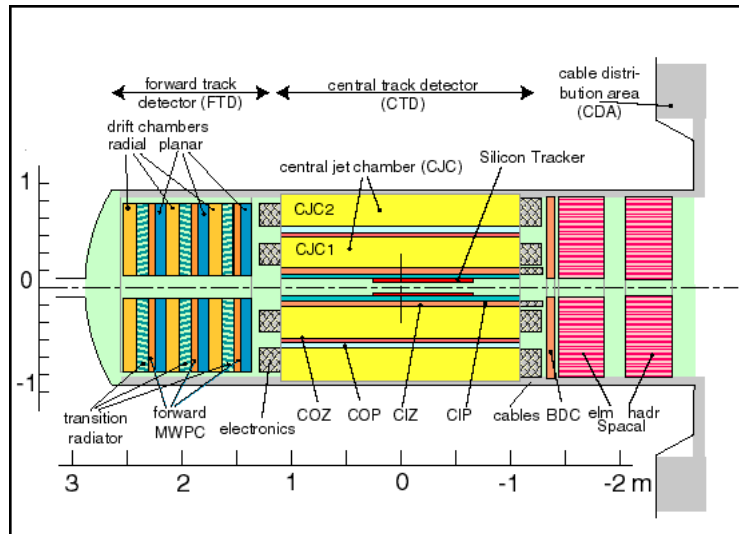


Figure 3.3: The tracking system (side view)

Central z-chambers

Between CST 1 and CJC 1 the CIZ is located. Together with the COZ, which lies between CJC 1 and CJC 2, it measures the z-component (parallel to the beam) of a track. The accuracy lies here at approx. $300 \mu\text{m}$. The wires run perpendicular to the beam and surround it concentric.

Central jet chambers

The wires of CJC 1 and CJC 2 are situated parallel to the beam and the drift chambers are radially tilted, causing the drift of ions to be more or less parallel to high momentum tracks in the magnetic field. The main task is here to make measurements in the $r\text{-}\phi$ plane. The accuracy here is approximately $170 \mu\text{m}$.

Forward tracking detector

This second part of the tracking system consists of a series of drift chambers and measures tracks with a polar angle $5^\circ < \theta < 25^\circ$.

3.2.2 Calorimeter

The task of the calorimeter is to measure the energy of the particles and to reconstruct the tracks of neutral hadrons and photons, which, electrically neutral, cannot be detected in the tracking system. Due to the interaction

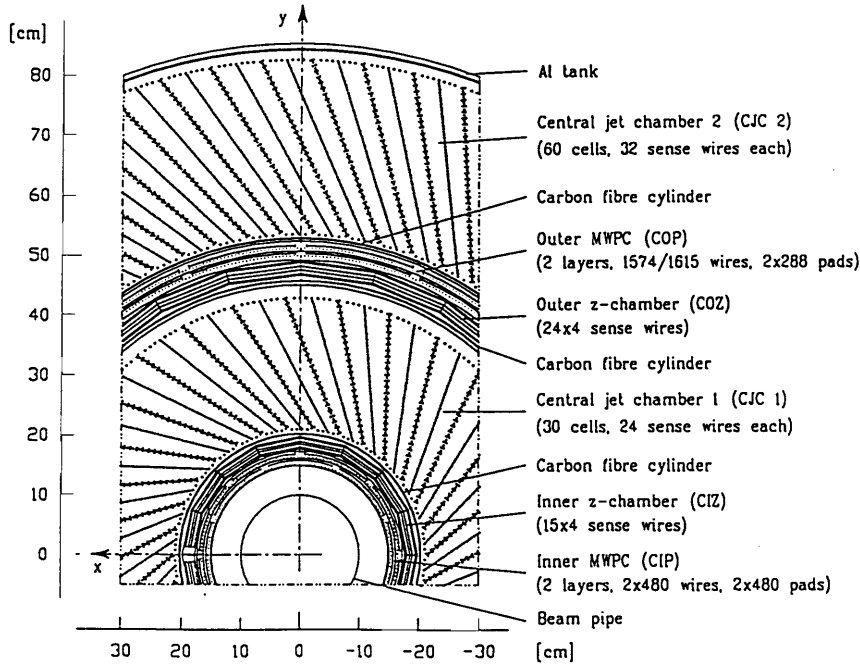


Figure 3.4: The tracking system (front view)

with the calorimeter material the entire energy of the particles is absorbed. Therefore all calorimeters of H1 contain absorbing (passive) layers and measuring (active) layers.

The calorimeters comprise an electromagnetic part and a hadronic part. In the electromagnetic part dense materials with high atomic number Z (e.g. lead) are used as absorbers. The particles then interact electromagnetically with the absorber material via *bremstrahlung* (electrons) or pair-production (photons). The result are showers of photons, electrons and positrons.

In the hadronic part, high absorbing masses (e.g. steel) are used. Here strong interaction takes place between the particles and the atoms of the absorber material. Again, showers are produced. They are broader than the electromagnetic ones making them easily distinguishable from the latter.

Since, in both cases the intensity of the showers is proportional to the absorbed energy, the energy can be measured by measuring the intensity of the showers.

The intensity is measured by the active material which registers and quantifies the showers.

The calorimeter consists of the liquid argon calorimeter (LAr) and spaghetti calorimeter (SpaCal).

The LAr has liquid argon as the active material and covers a polar angle of $4^\circ < \theta < 154^\circ$. A computer event display including the LAr is shown in fig. 3.5.

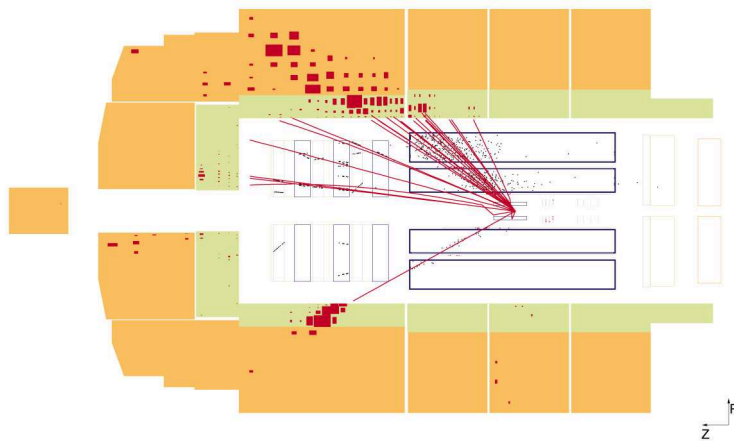


Figure 3.5: Event in the liquid argon calorimeter (side view)

The SpaCal uses scintillation fibers (therefore the name spaghetti) as active material and covers a polar angle of $153^\circ < \theta < 177.5^\circ$. Within these fibers the showers produce light proportional to their intensity. By total reflection this light can be transferred to a photomultiplier and there be transformed into an amplified electric signal. The SpaCal is shown in fig. 3.6.

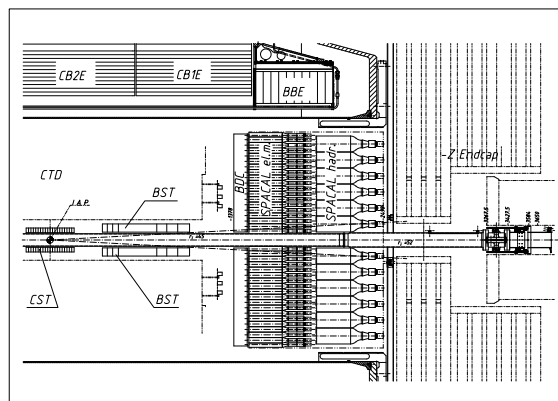


Figure 3.6: The spaghetti calorimeter (side view)

3.2.3 Muon detector

The muon energy cannot be completely absorbed in the calorimeter, because the muons simply leave the detector. This is why drift chambers are inserted between the layers of the iron yoke, whose task is simply to detect the passage of the muons.

3.3 Measurement of F_2

In the HERA the DIS kinematics are determined either by the scattered lepton, the hadronic final state or by a combination of both.

If the lepton is used, one speaks of the "electron method"². Here the kinematics are determined by the measured energy of the scattered lepton E'_e and its scattering angle relative to the z-axis³. With these quantities we get ([1]):

$$y_e = 1 - \frac{E'_e}{E_e} \sin^2 \left(\frac{\theta_e}{2} \right)$$

$$Q_e^2 = \frac{E_e'^2 \sin^2(\theta_e)}{1 - y_e} \quad (3.1)$$

The electron method becomes less accurate at low y ([1]). In this case the hadronic final state can be used to determine the kinematics. Using the so called " Σ method" ([1]) gives the following variables:

$$y_\Sigma = \frac{\sum_i (E_i - p_{i,z})}{\sum_i (E_i - p_{i,z}) + E'_e (1 - \cos(\theta_e))} \sin^2 \left(\frac{\theta_e}{2} \right)$$

$$Q_\Sigma^2 = \frac{E_e'^2 \sin^2(\theta_e)}{1 - y_\Sigma} \quad (3.2)$$

The relation (2.15) $Q^2 = sxy$ gives then - independent of the method - the Bjorken x .

The scattering angle of the lepton is measured by the tracking system of H1 and the lepton's energy by the SpaCal.

The hadronic final state is reconstructed by the tracking system, the SpaCal and the LAr.

²electron stands here for electron or positron

³The positive z axis in H1 is defined as the direction of the incident proton beam

According to eq. (2.59) the cross section is related to the structure functions via:

$$\frac{d^2\sigma}{dx dQ^2} = \frac{2\pi\alpha^2 Y_+}{xQ^4} \cdot \sigma_r \quad (3.3)$$

with the reduced cross section $\sigma_r = F_2(x, Q^2) - \frac{y^2}{Y_+} F_L(x, Q^2)$ and $Y_+ = 1 + (1 - y)^2$.

The longitudinal structure function F_L enters here multiplied with y^2 . Therefore its contribution is relevant only at high values of y and for most of the kinematic region one can assume $\sigma_r \approx F_2$.

Thus $F_2(x, Q^2)$ can be determined by measuring the cross section.

Some typical systematic errors in this measurement are ([1]):

- An electron beam energy calibration uncertainty of 1%.
- An electron polar angle uncertainty of 0.5 %.
- At lower angles $y \leq 0.1$ the kinematics cannot be determined by electron variables only and the hadronic part is used to determine the kinematics. Then, the uncertainty of the hadronic energy scale (2%) has to be taken into account.
- There is a background of $\approx 5\%$ due to photoproduction. Charged particles, such as π^- or e^- , produced in a photoproduction event can fake the scattered electron in the SpaCal.
- Beam background, caused by from electrons scattering from beam gas atoms and not from protons, results in an uncertainty of max. 3 %.
- The uncertainty of the luminosity is approx. 1.7%.

Chapter 4

Monte Carlo Event Generators

Simulations of high energy physics experiments have to take all particles entering and leaving a particle collision as well as their correct kinematics into account. Analytical calculations to describe all this would be too complicated to find solutions to them.

Perturbative methods such as the DGLAP equations take radiative corrections into account but usually do not treat the kinematics of the particles accurately. For inclusive measurements (as are total cross sections and F_2 measurements respectively) this poses no problem. But when it comes to exclusive measurements (as are p_t cross sections or jet measurements) the exact kinematics of the final state particles are crucial.

A Monte Carlo event generation has as an output all particles participating in an event as well as their proper kinematics. All stages are here taken into account, from the initial state with the colliding particles, exchange bosons and radiated gluons via the hard scattering to the final hadrons.

The basic principles of Monte Carlo event generators as well as the two generators used in this work will be introduced briefly in this chapter.

4.1 Monte Carlo Event Generators

The processes involved in an event are complex, including many physical aspects. Therefore, a description of a typical high energy physics process should simulate several subprocesses.

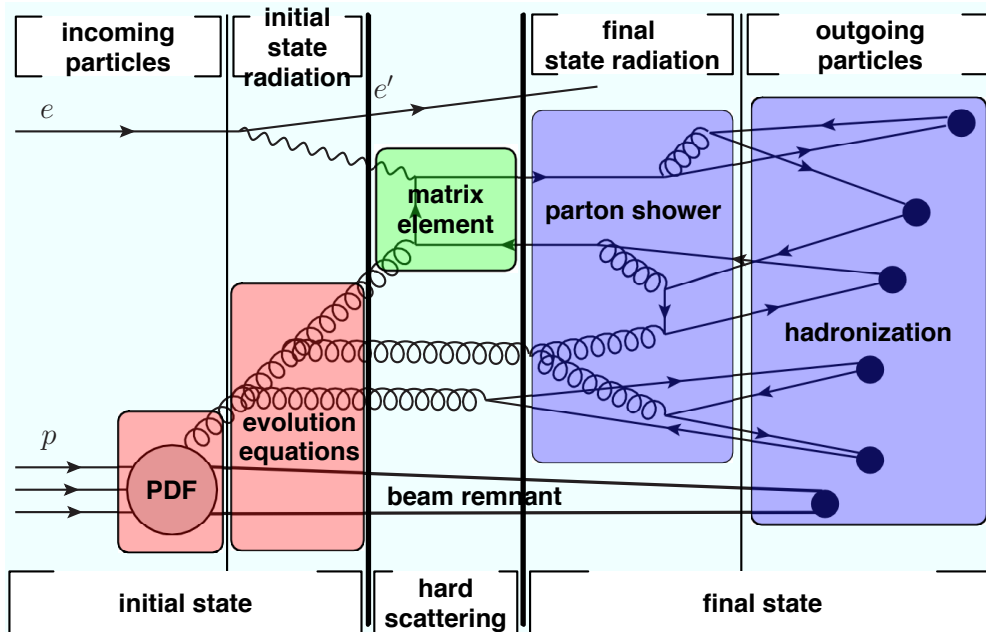


Figure 4.1: Schematic overview of the subprocesses in the generation of an event

The different stages that a full hadron level Monte Carlo event generator has to simulate and include are schematically shown in fig. 4.1 and consist of the following:

PDF: Initially two beam particles are flying towards each other. The momentum distributions of the incoming particles at a scale Q^2 are given by PDFs. In the case of the proton, for example, the PDFs (for valence quarks, sea quarks and gluons) contain all information about the momentum distribution of the partons in the proton at a scale Q^2 .

Initial state radiation: The perturbative part (accessible by perturbative QCD) happens here. A shower initiator parton from each beam starts a sequence of branchings. These emissions can, for example, be simulated by the Parton Shower Method (see sec. 4.2) from a lower cut off $Q_0^2 \approx 1$ GeV to an upper hardscale Q^2 given by the hard interaction. The Parton Shower Method makes it possible to evolve the PDFs perturbatively and to treat the kinematics of the particles accurately.

Below the value of Q_0^2 , α_s becomes too large to apply perturbative QCD and everything that happens below that scale is included in the non-perturbative part of the PDFs.

The transverse momentum distribution of the hard scattering is strongly influenced by the emission of gluons and photons in the initial (and final) state. The radiation of partons determines the kinematics of the final state particles.

Hard scattering: From each beam one of the particles enters here. The matrix element for the hard process is calculated. This is done according to the Feynman rules. The hard process cross section resulting from the matrix element can then, according the factorization theorem, be convoluted with the evolved PDFs.

Final state radiation and resonances: Short-lived resonances may be produced in the hard process which will then decay. Furthermore the outgoing partons will radiate just like the incoming did.

Hadronization: In the parton showers and the hard scattering usually many individual colored partons are produced. Due to QCD confinement these fragment in the final state to colorless hadrons so that no single quarks or gluons are directly observable. To describe this process phenomenological models are used.

Beam remnant: Only one of the partons of the proton interacts in DIS. But as it carries usually color, the rest, called beam remnant, is colored too. Thus, it is related to the rest of the final state.

4.2 Parton showers

The Monte Carlo event generators used in this work use the Parton Shower Method ([22, 24]), to describe the perturbative corrections introduced in ch. 2.3.

The Parton Shower Method models the radiation of partons by splittings $a \rightarrow bc$ of one parton a into two daughters b and c which continue to split. In the Monte Carlo event generators used in this work the evolution of the PDFs from the lower cut-off scale Q_0^2 on is described by the DGLAP equations (2.58). But other than in a mere evolution of PDFs, the Monte Carlo

event generator keeps track of the kinematics in each of the splittings. The splitting variable z (see ch. 2.3) describes how a 's momentum p_a is distributed over b 's momentum $p_b = zp_a$ and c 's momentum $p_c = (1 - z)p_a$. The value of z is determined by one of the four splitting functions of ch. 2.3 P_{qq} , P_{qg} , P_{gq} and P_{gg} . Each parton is characterized by a virtuality scale Q^2 at which the splitting happens.

In the initial state showers the virtuality Q^2 is increased, as the hard interaction is approached, until the upper scale $Q_{max,i}^2$, given by the hard interaction, is reached. The parton a entering the splitting $a \rightarrow bc$ and the parton b , which stays on the branch leading to the hard interaction, are spacelike, $-Q^2 = m^2 = E^2 - p^2 < 0$. With each splitting a gains more spacelike virtuality. This is in contrast to a pure DGLAP evolution, where, due to assumption of eq. (2.33), the virtuality gained at each splitting is neglected in the next interaction. Exactly this difference will affect the value of x at the hard scale as will become clear in ch. 9. The virtuality leads furthermore to a transverse momentum k_\perp of the initially collinear struck parton.

The parton c which is not taking part in the hard interaction compensates b 's increase in k_\perp and its virtuality is timelike. It will therefore produce a timelike shower in the final state.

In the final state the partons that leave the hard interaction have a timelike virtuality $-Q^2 = m^2 = E^2 - p^2 \geq 0$. From an upper scale $Q_{max,f}^2$ of the original parton the scale is evolved downwards until at some Q^2 a branching $a \rightarrow bc$ occurs. Each of these branchings decreases the virtuality until all final state partons are on shell and no further splittings are allowed.

To summarize: in the final state there are only timelike particles coming either from the hard scattering or from the timelike shower of the initial state radiation.

4.3 Hadronization

QCD perturbation theory is valid at short distances, where strong interaction can be treated as a weak perturbation. At long distances, the strong force increases to a level at which perturbative QCD breaks down. In this regime the colored partons turn into colorless hadrons. This is called hadronization and forms the last stage of the process (apart from decay that still might occur).

The hadronization mechanisms are not understood yet from first principles.

This is why some phenomenological models were developed to describe the process, such as the string model for hadronization and cluster hadronization. The first one is used by the Monte Carlo event generators employed in this work and so I will focus on the string model for hadronization.

In the Lund String Model two (color) charged objects are bound to each other by a color field which is squeezed together to form a narrow tube called *string*. This field is described by a Coulomb-like field plus a linear string potential. The color field has therefore the form:

$$V(r) = -\frac{4\alpha_s}{3r} + \kappa r \quad (4.1)$$

where r is the spacial distance between the objects and κ is the energy density of the string ($\kappa \approx 1$ GeV/fm).

The idea is now the following: due to the second term, the potential energy increases as the two objects move apart. Therefore more and more energy is stored in the string (as in a spring). This happens until it is energetically more favorable to break up the string and create a quark-antiquark pair. This process repeats until there is no sufficient energy left in the string to create new quarks and the hadrons formed by that process remain stable.

The two Monte Carlo event generators used in this work, PYTHIA and RAPGAP, use the Lund String Model for hadronization.

4.4 The Monte Carlo event generators of this work

In this work two Monte Carlo event generators, RAPGAP and PYTHIA, were used to calculate the structure function F_2 .

4.4.1 PYTHIA

PYTHIA 6.4 ([22]) is a multipurpose event generator that implements the DGLAP evolution equations and the Lund String Model for hadronization. It is well applicable for ep scattering and therefore for calculating F_2 . By means of a switch it gives the possibility to exclude parton shower or to include only initial state parton shower, only final state parton shower or the combination of initial and final state parton shower.

4.4.2 RAPGAP

RAPGAP 3.1 ([17]) is a multipurpose event generator as well. And like PYTHIA it implements the DGLAP evolution equations and the Lund String Model for hadronization and here as well the user has the possibility to choose between no parton shower, initial state parton shower, final state parton shower and initial + final state parton shower.

A major difference between the two generators is that in RAPGAP a method is implemented, that leaves the final kinematics Q^2 and x of the hard interacting parton unchanged by eventual parton showers. This method was proposed in ([3]) and later also implemented into RAPGAP and also in LEPTO([15]).

Chapter 5

Fitting

5.1 General formalism

In ch. 2 the parton density functions (PDFs) $f_i(x, Q^2)$ were introduced as a parton number density. In leading order α_s calculations the PDFs are Q^2 -dependent and $f_i(x, Q^2)dx$ gives the probability of finding a parton of flavour i with momentum fraction x at a certain Q^2 .

From the DGLAP equations (2.58) the evolution of the PDFs with $\ln(Q^2)$ is known, but not their x -dependence at the starting scale of the evolution. The analytic form of a PDF at some Q^2 is not calculable from first principles, but its evolution with Q^2 can be calculated.

Therefore the PDF at a starting scale Q_0^2 is guessed in a parametrized form. The value of the starting scale Q_0^2 can be arbitrary but should be sufficiently large to assume $\alpha_s(Q^2)$ to be small enough in order to make perturbative calculations applicable.

The next step is evolving the PDFs with e.g. the DGLAP equations (2.58) to some other value of Q^2 . The PDFs can then be used to calculate the structure functions (see eqns. (2.39) and (2.43)). The parameters of the PDFs at the starting scale Q_0^2 can then be adjusted such that the calculations describe the data. This procedure is called fitting. The quality of the fit is usually determined by the quantity χ^2 which is described in ch. 6.

The parametrization of the PDFs at the starting scale Q_0^2 usually has the form:

$$xf(x, Q_0^2) = A_f x^{-B_f} (1-x)^{C_f} P(x, f) \quad (5.1)$$

with $P(x, f)$ being a polynomial in x or \sqrt{x} .

The two terms x^{-B_f} and $(1-x)^{C_f}$ describe the behaviour of the PDFs at

$x \rightarrow 0$ and $x \rightarrow 1$.

B_f can be guessed from Regge theory ([8]).

The $(x \rightarrow 1)$ -behaviour is described by $(1-x)^{C_f}$ for $C_f \neq 0$. It reflects the fact that the PDFs should tend to zero when $x \rightarrow 1$, because then there is no momentum left for the other partons than the struck one. The other partons therefore become "spectators". The so called counting rules give a rough estimate for the value of C_f . The prediction is $C_f \approx 2n_s - 1$, where n_s is the minimum number of spectators, e.g. for the 3 valence quarks (uds) $n_s = 2$ and so $C_f \approx 3$ ([8]).

5.2 Minimization

To find the minimum of a function $F(p)$ the requirement is:

$$F(p^*) < F(p) \quad \forall p \quad (5.2)$$

In practice the conditions used are often based on the derivatives. But they require smooth functions, meaning differentiable and continuous. The conditions then are:

- **1-dimensional case**

$$\frac{dF(p)}{dp} \Big|_{p=p^*} = 0 \quad \wedge \quad \frac{d^2F(p)}{dp^2} \Big|_{p=p^*} > 0 \quad (5.3)$$

- **n-dimensional case**

here one uses the n-dimensional gradient $\nabla F(p)$ and the $n \times n$ Hessian matrix

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2 F(p)}{\partial p_1^2} & \frac{\partial^2 F(p)}{\partial p_1 \partial p_2} & \cdots & \frac{\partial^2 F(p)}{\partial p_1 \partial p_n} \\ \frac{\partial^2 F(p)}{\partial p_2 \partial p_1} & \frac{\partial^2 F(p)}{\partial p_2^2} & \cdots & \frac{\partial^2 F(p)}{\partial p_2 \partial p_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F(p)}{\partial p_n \partial p_1} & \frac{\partial^2 F(p)}{\partial p_n \partial p_2} & \cdots & \frac{\partial^2 F(p)}{\partial p_n^2} \end{pmatrix}. \quad (5.4)$$

The conditions for a minimum then are:

$$\nabla F(p) = \vec{0} \quad \wedge \quad \mathbf{H} \text{ positive definite} \quad (5.5)$$

5.3 The SIMPLEX method

In this work the MINUIT program ([16]) was used to perform the fits. MINUIT offers different methods for minimization. The one used here was the SIMPLEX method by *Nelder* and *Mead*¹ as now briefly described.

Imagine a function containing n parameters p_i that has to be fitted to e.g. some data. Every combination of the parameters p_i can be understood as a point $\vec{x}_j = (p_{1,j}, p_{2,j}, \dots, p_{n,j})$ in an n -dimensional parameter space. For the best parameter set given by some point \vec{x}_{min} the condition (5.2) or some of its alternative forms must be fulfilled. The SIMPLEX method searches in the n -dimensional parameter space for a minimum by comparing points \vec{x}_i with each other and is therefore a direct search algorithm. In other words: it is not using derivatives to find the minimum but implements the condition (5.2). Therefore it is well suited to optimize functions that are not differentiable or are discontinuous and where it is difficult to find the derivatives numerically with the clear advantage, that it is a fast method.

This makes it a good choice when fitting Monte Carlo predictions to data, because the former ones are usually not smooth functions and show statistical fluctuations. Not smooth means having no differentiable or discontinuous points and thus making it impossible to determine a derivative at all.

However the SIMPLEX method has a major drawback: it has difficulties in finding global minima. It is sensitive to mistaking local minima for global ones. One has to interpret the results of a fit, although converging, therefore very carefully.

5.3.1 The SIMPLEX algorithm

To optimize n parameters the SIMPLEX algorithm ([4]) examines in n -dimensional parameter space at every cycle $n + 1$ points \vec{x}_1 to \vec{x}_{n+1} , where $\vec{x}_i = (p_i^{(1)}, \dots, p_i^{(n)})$ gives a point in parameter space standing for one set of parameter values.

$n + 1$ points are needed because a simplex (polyhedron) is spanned by the points \vec{x}_i in n -dimensional parameter space (e.g. a triangle in 2D and a tetrahedron in 3D). SIMPLEX works in cycles and moves in every cycle the simplex away from the point \vec{x}_{n+1} standing for the parameter set describing the data worst for all $n + 1$ points.

Every cycle of the algorithm then contains the following steps (an example in 2-dim. parameter space is shown in fig. 5.1):

¹the method is not to be confused with the *simplex algorithm* of linear programming

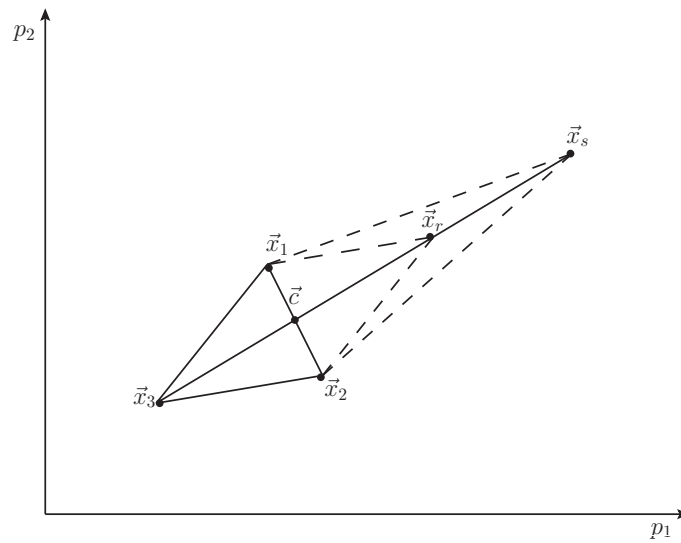


Figure 5.1: SIMPLEX for 2 parameters: the three points in 2-D parameter space \vec{x}_1 , \vec{x}_2 and \vec{x}_3 form a simplex with the center of mass (the centroid) \vec{c} . Furthermore two test points \vec{x}_r and \vec{x}_s are plotted.

1. Order the $n + 1$ points such that

$$F(\vec{x}_1) < F(\vec{x}_2) < \dots < F(\vec{x}_{n+1})$$

where $F(\vec{x})$ is the function to be optimized.

2. Calculate the center of mass of the n best points with lowest value $F(\vec{x}_i)$ - the so called centroid,

$$\vec{c} = \frac{1}{n} \sum_{i=1}^n \vec{x}_i$$

3. Create a test point \vec{x}_r by reflecting the point with the highest functional value (the worst point) \vec{x}_{n+1} through the centroid \vec{c} via

$$\vec{x}_r = \vec{c} + \alpha(\vec{c} - \vec{x}_{n+1})$$

with the constant factor $\alpha > 0$.

4. The next step depends then on the functional value $F(\vec{x}_r)$:

- (a) $F(\vec{x}_1) \leq F(\vec{x}_r) \leq F(\vec{x}_n)$

\vec{x}_r is not better than \vec{x}_1 and not worse than \vec{x}_n . Therefore: replace \vec{x}_{n+1} by \vec{x}_r and go to step 1).

(b) $F(\vec{x}_r) \leq F(\vec{x}_1)$

The test point is the new best point. So continue in the search direction and create a second test point $\vec{x}_s = \vec{c} + \beta(\vec{x}_r - \vec{c})$ with $\beta > 1$. If now

- $F(\vec{x}_s) < F(\vec{x}_r)$, replace \vec{x}_{n+1} by \vec{x}_s and go to step 1)
- $F(\vec{x}_s) \geq F(\vec{x}_r)$, replace \vec{x}_{n+1} by \vec{x}_r and go to step 1)

(c) $F(\vec{x}_r) > F(\vec{x}_n)$ If $F(\vec{x}_r) < F(\vec{x}_{n+1})$ replace the worst point \vec{x}_{n+1} by the reflected point \vec{x}_r else discard \vec{x}_r . Then make the simplex smaller. Therefore create a new test point $\vec{x}_s = \vec{c} - \gamma(\vec{c} - \vec{x}_{n+1})$ with $0 < \gamma < 1$. If now

- $F(\vec{x}_s) < F(\vec{x}_{n+1})$, replace \vec{x}_{n+1} by \vec{x}_s and go to step 1).
- $F(\vec{x}_s) \geq F(\vec{x}_{n+1})$ replace all points but the best point \vec{x}_1 by $\vec{x}_j = \vec{x}_1 + \delta(\vec{x}_j - \vec{x}_1)$ with $0 < \delta < 1$ and go to step 1).

The criterion for convergence in the MINUIT version is given by either a maximum number of iterations or an estimated distance to the minimum (EDM) less than a set value.

Chapter 6

Treatment of correlated systematic uncertainties

To see how good a prediction $\{t_i\}$ describes a set of measured data $\{d_i\}$ the deviation of the t_i from the d_i is compared to the uncertainties of both, prediction and data.

On the experimental side there are for each data point statistical uncertainties σ_i^{stat} and systematic uncertainties¹ σ_i^{yst} . But also the prediction can have uncertainties, e.g. the statistical fluctuation of a Monte Carlo prediction σ_i^{theo} or systematic uncertainties such as model uncertainties or uncertainties from the input.

Uncorrelated and correlated uncertainties have to be differentiated as well. Two uncertainties are said to be uncorrelated if the value of one uncertainty is independent of the value of the other. e.g. the uncertainty of a measurement of the height of Fujiyama and the uncertainty of a measurement of my body temperature are uncorrelated.

On the other hand two uncertainties are said to be correlated if the value of one uncertainty depends on the value of the other uncertainty. An example would be the uncertainty of a measured room temperature and the uncertainty of reading a ruler in that room. As the ruler expands due to temperature, the total reading accuracy depends on the uncertainty of temperature and both uncertainties are therefore correlated.

In the case of experiments, correlated uncertainties mean that fluctuations in the single bins due to these uncertainties are not independent. For instance,

¹A systematic uncertainty can be correlated or uncorrelated. An example for a correlated systematic uncertainty is the luminosity uncertainty. Here all data points are affected in a similar manner. An example for an uncorrelated systematic uncertainty is the measurement of the transverse energy of a particle. Here every data point is affected individually.

if a data point in one bin is influenced in a certain way by a correlated uncertainty, then all others are influenced in the same way. If one looks in contrast at the uncorrelated uncertainties, the influence of it on one bin can be totally different from the influence on other bins.

Some typical correlated systematic uncertainties of an F_2 -measurement at HERA are given in sec. 3.3:

In a classical ansatz the quantity χ_0^2 is used to determine the quality of a fit, e.g. of experimental data to their prediction.

$$\chi_0^2 = \sum_{i=1}^{N_{dat}} \frac{(d_i - t_i)^2}{(\sigma_i^{dat})^2 + (\sigma_i^{theo})^2} \quad (6.1)$$

where N_{dat} is the number of data points used, d_i is the i -th data point and t_i is the corresponding i -th point of the prediction. $(\sigma_i^{dat})^2 = (\sigma_i^{stat})^2 + \sum (\sigma_i^{syst})^2$ is the quadratic sum of the statistical experimental uncertainty σ_i^{stat} and the point-to-point systematic uncertainties σ_{ki}^{syst} which can, assuming to be uncorrelated, be added in quadrature.

As χ^2 is the sum over the ratios of the quadratic deviation of prediction and data over the standard deviation, one can use

$$\chi^2 / (N_{dat} - n_{par}) = \chi^2 / ndf, \quad (6.2)$$

where n_{par} is the number of fitting parameters and $N_{dat} - n_{par} =: ndf$ is the number of degrees of freedom, as a good measure for the quality of the fit.

If, as presented in this analysis, a Monte Carlo generator is used to produce predictions the total statistical uncertainty of the prediction σ_i^{theo} must be taken into account as well, as is done in (6.1).

Determining χ_0^2 is a convenient method in the search for an optimal fit. But as there are often many correlated systematic uncertainties present in a measurement these need to be properly taken into account.

Fortunately methods were developed to assess the uncertainties properly. In this analysis methods were used to include a full treatment of correlated systematic uncertainties.

To compare the impact of correlated systematic uncertainties on the fitting procedure the classical χ_0^2 as given in 6.1 was determined as well.

In the following three chapters the general ansatz to take correlated systematic uncertainties into account as well as two realizations will be described. The first one follows an ansatz proposed by CTEQ ([7]) and will therefore be

called CTEQ-like in the following. The second one is proposed by Pascaud et al. ([6]) and will be referred to as Pascaud-like.

6.1 General ansatz

Most experiments provide now detailed information on the systematic uncertainties of their measurements. There are for each data point a statistical experimental uncertainty σ_i^{dat} , an uncorrelated systematic uncertainty u_i^{syst} and n_{sys} correlated systematic uncertainties $\{\beta_{i1}^{syst}, \beta_{i2}^{syst}, \dots, \beta_{in_{sys}}^{syst}\}$ (in the interest of readability I will skip the superscript "syst" for u and β).

Including correlated systematic uncertainties properly into χ^2 leads to the formula:

$$\chi^2(\{a\}, \{r\}) = \sum_{i=1}^N \left(\frac{d_i - t_i - \sum_{k=1}^{n_{sys}} \beta_{ik} r'_k}{\alpha_i} \right)^2 + \sum_{k=1}^{n_{sys}} r'_k{}^2 \quad (6.3)$$

where $\alpha_i^2 = (\sigma_i^{dat})^2 + u_i^2$ is the combined uncorrelated uncertainty.

The t_i depend on the parameters $\{a\}$ and thus on χ^2 . Furthermore χ^2 depends on the random parameters $\{r\}$ associated with the correlated systematic uncertainties.

Following [23] a brief derivation of the formula eq. (6.3) is sketched.

Assuming there is a set of true values of the observable $\{O_i\}$ and also a prediction, that describes these "true" values perfectly so that $\{O_i\} = \{t_i\}$.

The measurements $\{d_i\}$ will in general deviate from the set of "true" values $\{t_i\}$. The deviation will be caused by various kinds of uncertainties. For the i -th measurement there are the uncorrelated uncertainties, summed up in quadrature yielding α_i , and the n_{sys} correlated systematic uncertainties $\{\beta_{ik}\}$. Thus the measured value d_i equates as:

$$d_i = t_i + r_i \alpha_i + \sum_{k=1}^{n_{sys}} r'_k \beta_{ik} \quad (6.4)$$

where the r_i, r'_k express the individual shifts of the data points by the uncertainties (e.g. $r_i = 1$ means a shift of the i -th datapoint by 1σ). They should be of the order of 1 and follow a Gaussian distribution with width 1 ([19]), e.g.

$$p(r) \propto e^{-r^2/2} \quad (6.5)$$

The probability distribution of the measurements therefore is,

$$dP = \left(\int \prod_{i=1}^N p(r_i) dr_i \prod_{k=1}^{n_{sys}} p(r'_k) dr'_k \prod_{i=1}^N \delta(d_i - t_i - r_i \alpha_i - \sum_{k=1}^{n_{sys}} r'_k \beta_{ik}) \right) d^N d \quad (6.6)$$

with the differential $d^N d = d d_1 \cdot d d_2 \cdot \dots \cdot d d_N$

First integrate over the r_i by using the delta distributions and (6.5):

$$\int \prod_{i=1}^N p(r_i) dr_i \prod_{i=1}^N \delta(d_i - t_i - r_i \alpha_i - \sum_{k=1}^{n_{sys}} r'_k \beta_{ik}) \propto \prod_i e^{\frac{1}{2} \left(\frac{d_i - t_i - \sum_{k=1}^{n_{sys}} \beta_{ik} r'_k}{\alpha_i} \right)^2}$$

Using (6.5) again in the integral over the r'_k gives

$$\int dr'_k \prod_{k=1}^{n_{sys}} p(r'_k) \propto \int dr'_k e^{\frac{1}{2} \sum_{k=1}^{n_{sys}} r_k'^2}$$

Putting both results together gives:

$$dP = \left(\int \prod_{k=1}^{n_{sys}} dr'_k C_1 e^{-\chi^2/2} \right) d^N d \quad (6.7)$$

with the normalization constant C_1 and

$$\chi^2(\{a\}, \{r\}) = \sum_{i=1}^N \left(\frac{d_i - t_i - \sum_{k=1}^{n_{sys}} \beta_{ik} r'_k}{\alpha_i} \right)^2 + \sum_{k=1}^{n_{sys}} r_k'^2 \quad (6.8)$$

This is the expression of (6.3) which, in order to find a best fit, has to be minimized with respect to $\{a\}$ (the ones of actual interest) and $\{r\}$. The total number of fitting parameters is given by the number of the $\{a\}$ and $\{r\}$ and can become quite large (> 50). This can take a lot of computing time. Thus, it is convenient to try to reduce the task to fitting only the $\{a\}$.

To do so, one first expands the quadratic term and rewrites χ^2 :

$$\begin{aligned} \chi^2 &= \sum_{i=1}^N \frac{1}{\alpha_i^2} \left((d_i - t_i)^2 + \left(\sum_{k=1}^{n_{sys}} \beta_{ik} r'_k \right)^2 - 2(d_i - t_i) \sum_{k=1}^{n_{sys}} \beta_{ik} r'_k \right) + \sum_{k=1}^{n_{sys}} r_k'^2 \\ &= \sum_{i=1}^N \frac{(d_i - t_i)^2}{\alpha_i^2} + \sum_{\tilde{k}, k=1}^{n_{sys}} \left(\sum_{i=1}^N \frac{\beta_{ik} \beta_{i\tilde{k}}}{\alpha_i^2} + \delta_{k\tilde{k}} \right) r'_k r'_{\tilde{k}} - \sum_{k=1}^{n_{sys}} 2 \frac{\beta_{ik} (d_i - t_i)}{\alpha_i^2} r'_k \quad (6.9) \\ &= \sum_{i=1}^N \frac{(d_i - t_i)^2}{\alpha_i^2} + \sum_{\tilde{k}, k=1}^{n_{sys}} A_{k\tilde{k}} r'_k r'_{\tilde{k}} - \sum_{k=1}^{n_{sys}} 2 B_k r'_k \end{aligned}$$

where

$$B_k = \sum_{i=1}^{N_{dat}} \frac{\beta_{ik}(d_i - t_i)}{\alpha_i^2} \quad \text{and} \quad A_{k\tilde{k}} = \delta_{k\tilde{k}} + \sum_{i=1}^{N_{dat}} \frac{\beta_{ik}\beta_{i\tilde{k}}}{\alpha_i^2}. \quad (6.10)$$

χ^2 is a quadratic polynomial of the r'_k . Demanding

$$\frac{\partial \chi^2}{\partial r'_k} = 0, \quad \forall \in [1, n_{sys}]$$

it can be minimized analytically with respect to the r'_k . This minimization leads to the best estimate for the systematic uncertainties

$$r'_k(\{a\}) = \sum_{k'=1}^{n_{sys}} (A^{-1})_{kk'} B_{k'} \quad (6.11)$$

To get an r'_k -independent expression for χ^2 one inserts this result into (6.9)

$$\begin{aligned} \chi^2 = & \sum_{i=1}^{N_{dat}} \frac{(d_i - t_i)^2}{\alpha_i^2} - \sum_{k,k'=1}^{n_{sys}} 2B_k (A^{-1})_{kk'} B_{k'} \\ & + \sum_{k\tilde{k}}^{n_{sys}} A_{k\tilde{k}} \left(\sum_{k''}^{n_{sys}} (A^{-1})_{\tilde{k}k''} B_{k''} \right) \left(\sum_{k'}^{n_{sys}} (A^{-1})_{kk'} B_{k'} \right) \end{aligned} \quad (6.12)$$

The last of the three terms can be transformed

$$\begin{aligned} & \sum_{k\tilde{k}}^{n_{sys}} A_{k\tilde{k}} \left(\sum_{k''}^{n_{sys}} (A^{-1})_{\tilde{k}k''} B_{k''} \right) \left(\sum_{k'}^{n_{sys}} (A^{-1})_{kk'} B_{k'} \right) \\ & = \sum_{kk''}^{n_{sys}} \left(\sum_{\tilde{k}}^{n_{sys}} A_{k\tilde{k}} (A^{-1})_{\tilde{k}k''} \right) B_{k''} \sum_{k'}^{n_{sys}} (A^{-1})_{kk'} B_{k'} \\ & = \sum_{kk''}^{n_{sys}} \delta_{kk''} B_{k''} \sum_{k'}^{n_{sys}} (A^{-1})_{kk'} B_{k'} \\ & = \sum_{kk'}^{n_{sys}} B_k (A^{-1})_{kk'} B_{k'} \end{aligned}$$

In the second step I used the fact $(A \cdot A^{-1})_{kk''} = \mathbf{1}_{kk''} = \delta_{kk''}$.

As a final expression one gets now:

$$\chi^2 = \sum_{i=1}^{N_{dat}} \frac{(d_i - t_i)^2}{\alpha_i^2} - \sum_{k,k'=1}^{n_{sys}} B_k (A^{-1})_{kk'} B_{k'} \quad (6.13)$$

This is a simplified version of (6.3) and the the problem is reduced to the irreducible task of fitting the parameters $\{a\}$ only.

If a Monte Carlo event generator is used to make the predictions $\{t_i\}$ the statistical uncertainty of the Monte Carlo event generator has to be taken into account as well (as in (6.1)). This is done by including this uncertainty in the uncorrelated uncertainties and redefining

$$\alpha_i^2 = (\sigma_i^{dat})^2 + u_i^2 + (\sigma_i^{theo})^2. \quad (6.14)$$

In a nutshell, the difference of (6.3) from the simple form (6.1) is as follows: in (6.3) the correlated systematic uncertainties β_{ik} are not simply added in quadrature (which would require them to be uncorrelated and as done in (6.1)) but treated separately from the uncorrelated uncertainties. The latter ones can simply be added in quadrature and are included in α_i . The idea is, that there are fluctuations $\sum_{k=1}^{n_{sys}} r_k \beta_{ik}$ due to systematics.

The two following chapters will now treat two main realizations of the ideas developed in this chapter. The CTEQ-like method assumes, that the systematics shift the data points while the Pascaud-like ansatz follows the idea, that the systematics shift the theory values.

6.2 CTEQ-like treatment of correlated systematic uncertainties

In the CTEQ-like ansatz it is assumed that the experimental data are shifted by the systematics in the following way:

$$d_i = t_i + r_i \alpha_i + \sum_{k=1}^{n_{sys}} r'_k \beta_{ik} = t_i + r_i \alpha_i + \sum_{k=1}^{n_{sys}} r'_k (\delta_{ik} d_i) \quad (6.15)$$

where $\delta_{ik} = \frac{\beta_{ik}}{d_i}$ is the relative correlated systematic uncertainty.

Using $\beta_{ik} = \delta_{ik} d_i$ gives together with (6.3) the formula for χ^2 :

$$\chi^2(\{a\}, \{r\}) = \sum_{i=1}^N \left(\frac{d_i - t_i - \sum_{k=1}^{n_{sys}} (\delta_{ik} d_i) r'_k}{\alpha_i} \right)^2 + \sum_{k=1}^{n_{sys}} r'_k{}^2 \quad (6.16)$$

or in the simplified form (6.13)

$$\chi^2 = \sum_{i=1}^{N_{dat}} \frac{(d_i - t_i)^2}{\alpha_i^2} - \sum_{k,k'=1}^{n_{sys}} B_k (A^{-1})_{kk'} B_{k'}$$

with

$$r'_k(\{a\}) = \sum_{k'=1}^{n_{sys}} (A^{-1})_{kk'} B_{k'} \quad (6.17)$$

and

$$B_k = \sum_{i=1}^{N_{dat}} \frac{\delta_{ik} d_i (d_i - t_i)}{\alpha_i^2} \quad \text{and} \quad A_{kk'} = \delta_{kk'} + \sum_{i=1}^{N_{dat}} \frac{\delta_{ik} \delta_{ik'} d_i^2}{\alpha_i^2} \quad (6.18)$$

6.3 Pascaud-like treatment of correlated systematic uncertainties

In the Pascaud-like ansatz it is assumed that the theory values are shifted by the systematics rather than by the experimental values (as in the CTEQ-like ansatz). This assumes the relation:

$$d_i = t_i + r_i \alpha_i + \sum_{k=1}^{n_{sys}} r'_k \beta_{ik}^p = t_i + r_i \alpha_i + \sum_{k=1}^{n_{sys}} r'_k (\delta_{ik} t_i) \quad (6.19)$$

where $\delta_{ik} = \frac{\beta_{ik}}{d_i}$ is again the relative correlated systematic uncertainty.

But when now assuming that $\beta_{ik}^p = \delta_{ik} t_i = \frac{\beta_{ik}}{d_i} t_i$ (therefore the two β are here distinguished by the superscript p). This yields together with (6.3) the formula for χ^2 as used in the Pascaud-like ansatz:

$$\begin{aligned} \chi^2(\{a\}, \{r\}) &= \sum_{i=1}^N \left(\frac{d_i - t_i - \sum_{k=1}^{n_{sys}} (\delta_{ik} d_i) r'_k}{\alpha_i} \right)^2 + \sum_{k=1}^{n_{sys}} r_k'^2 \\ &= \sum_{i=1}^N \left(\frac{d_i - t_i (1 + \sum_{k=1}^{n_{sys}} \delta_{ik} r'_k)}{\alpha_i} \right)^2 + \sum_{k=1}^{n_{sys}} r_k'^2 \end{aligned} \quad (6.20)$$

or again in the simplified form (6.13)

$$\chi^2 = \sum_{i=1}^{N_{dat}} \frac{(d_i - t_i)^2}{\alpha_i^2} - \sum_{k,k'=1}^{n_{sys}} B_k (A^{-1})_{kk'} B_{k'}$$

where

$$r'_k(\{a\}) = \sum_{k'=1}^{n_{sys}} (A^{-1})_{kk'} B_{k'} \quad (6.21)$$

and

$$B_k = \sum_{i=1}^{N_{dat}} \frac{\delta_{ik} t_i (d_i - t_i)}{\alpha_i^2} \quad \text{and} \quad A_{kk'} = \delta_{kk'} + \sum_{i=1}^{N_{dat}} \frac{\delta_{ik} \delta_{ik'} t_i^2}{\alpha_i^2} \quad (6.22)$$

Chapter 7

Data, software and Monte Carlo tuning

7.1 Dataset

The parton densities are related to the structure function $F_2(x, Q^2)$ by eqns. (2.39) and (2.43). Therefore parameterizations for the PDFs can be obtained by fits of predictions for $F_2(x, Q^2)$ to F_2 -data from DIS measurements. The convention is to speak of data, when the data come from real measurements, in contrast to Monte Carlo predictions. I will follow this convention in the text and make explicit distinctions in the denomination only when necessary.

In this thesis the cross section measurement from [1] is used. This paper describes cross section measurements for the neutral current process $e^+p \rightarrow e^+X$. The kinematic region covered here is $1.5 \leq Q^2 \leq 150 \text{ GeV}^2$ and $3 \cdot 10^{-5} \leq x \leq 0.2$. In this thesis 104 data points with $5 \leq Q^2 \leq 150 \text{ GeV}^2$ were used. The measurements were obtained from data taken in the years 1996 and 1997. At that time¹ the positron energy was $E_e = 27.6 \text{ GeV}$ and the proton energy $E_p = 820 \text{ GeV}$. This corresponds to a center of mass energy of $\sqrt{s} = 300.9 \text{ GeV}$. More recent results were published in the meantime but were not available when this thesis was being written.

¹From 1998 on the protons energy was raised to $E_p = 920 \text{ GeV}$. Compare ch. 3

7.2 Tools

7.2.1 HZTOOL

To compare data to Monte Carlo predictions the HZTOOL package is used. HZTOOL ([5]) is a collection of Fortran routines, that can be interfaced with a Monte Carlo event generator. HZTOOL routines can be associated to a publication of data and contains then the data of the publication. When the routine is executed it writes these data into a histogram, reads the output of the interfaced Monte Carlo event generator and writes the generator output into a properly normalized² histogram. In the form of histograms the data and the predictions can be analyzed, compared to each other and/or plotted in a convenient way.

In an HZTOOL routine a selection of events can be made by applying *cuts* (e.g. events lying in some x -range given by the cuts x_{min} and x_{max} are selected). Only these events are then written into a histogram. Cuts are often defined by the limitations of the detector.

The typical structure of a HZTOOL routine is the following:

1. **Initialization**

The data is declared in the form of arrays, and the histograms for the data and the Monte Carlo predictions are declared³, but yet to be filled.

2. **Event generation**

The Monte Carlo event generator then produces a number of events, set by the user, and the result of every single event is written into the corresponding prediction histogram. Cuts can be applied here, such that only events satisfying the cuts are filled into a histogram.

3. **Termination**

The data histograms are then filled, and the prediction histograms filled in step 2 normalized.

The HZTOOL-routine used in this case was *hz00181* ([5]) and refers to [1].

7.2.2 The routine *errortreat*

Part of this work was to implement a routine that calculates χ^2 with full treatment of the systematic errors. It was realized in the Fortran routine

²Usually as described in ch. 7.3

³The convention is that data histograms are named by h_n where n is a number, while the according prediction is named without the underscore, hn .

errortreat(*chi2*,*ichi2*,*theo*,*etheo*,*dat*,*dtot*,*dst*,*dsyu*,*beta*). It calculates the χ^2 regarding to one of the three methods described in ch. 6 - the classical ansatz, the Pascaud-like method or the CTEQ-like method. It thereby offers besides the classical ansatz the possibility to determine the χ^2 with full treatment of systematic errors. The choice of the method to calculate χ^2 is done by the user via a switch.

The routine expects the following arguments:

chi2: this is simply the quantity χ^2 which is to be calculated and is the output of the routine.

ichi2: this is the switch, that determines the method of how χ^2 is calculated. 1: Pascaud-like, 2: CTEQ-like, 3: classical ansatz

theo: the predictions given in form of an array. The order must be the same as the order of the according data (see *dat*).

etheo: the statistical errors of the predictions given in form of an array.

dat: the data given in form of an array.

dtot: the total errors of the data given in form of an array. The definition of that error depends on the user. Here it was assumed to be the quadratic sum of the statistical data error and the uncorrelated systematic errors (s.f. ch. 6).

dst: the statistical errors of the data given in form of an array.

dsyu: the uncorrelated errors of the data given in form of an array.

beta: the correlated systematic errors of the data given in form of a matrix *beta*(*i*,*j*). Thereby *i* gives the number of the data point and *j* the number of the correlated systematic error. This is necessary as there are usually a few correlated systematic errors for each data point.

All errors have to be given as absolute errors.

The routine *errortreat* is basically independent of how the data is presented (as an HZTOOL routine or a text file etc.). The data in this specific case was not given by a HZTOOL routine, but in form of a simple text file containing the data in lines and columns.

The text file is then read in by a routine.

In the existing framework a routine *f2_fit* did the job. It reads in the data including errors from the text file and the predictions including errors out of

the histogram created by *hz00181*.

Nevertheless *f2_fit* was modified such that it prepares the data in the format demanded by the routine *errortreat*. *f2_fit* then calls, after an optional application of cuts, *errortreat*. The calculated χ^2 is then passed on by *f2_fit* to the fitting routine MINUIT.

7.3 Histograms, bins and the determination of cross sections

Most high energy experiments and simulations involve counting the events. In a measurement or physics simulation of an Monte Carlo event generator the number of events N and the cross section are related by

$$N = \mathcal{L}\sigma \quad (7.1)$$

where σ is the total cross section, \mathcal{L} is the luminosity (number of beam particles per second per unit area).

The number of events is usually plotted in histograms. The abscissa of a histogram is divided into several segments, each standing for a certain range of the abscissa's variable. These ranges are called *bins*. The ordinate of a histogram gives the total number of events or the cross section.

The relation of a differential cross section, say the double differential cross section $\frac{d^2\sigma}{dQ^2 dx}$, to the number of events ΔN in the bin of width $\Delta Q^2 \Delta x$ is given by

$$\Delta Q^2 \Delta x \frac{d^2\sigma}{dQ^2 dx} = \frac{\Delta N}{\mathcal{L}} \quad (7.2)$$

To give a simple example: at a luminosity of $\mathcal{L} = 100 \frac{\text{GeV}}{\text{pb}}$ and with $N_i = 10^5$ events in the p_t -range of 20-30 GeV regarding to (7.2) the differential cross section would be :

$$\frac{d\sigma}{dp_t} \approx \frac{N_i}{\Delta p_t \mathcal{L}} = \frac{10^5}{100 \frac{\text{GeV}}{\text{pb}} \cdot 10 \text{ GeV}} = 100 \text{ pb}$$

where $\Delta p_t = 30 - 20 \text{ GeV} = 10 \text{ GeV}$ is the bin width of the p_t bin.

This means that to obtain differential cross sections from the output of a Monte Carlo event generator the number of events has to be normalized bin per bin to the luminosity \mathcal{L} and the bin width of the variable of interest (in the example Δp_t).

7.4 Tuning the Monte Carlo generators RAPGAP and PYTHIA

In this analysis the two Monte Carlo generators RAPGAP and PYTHIA were used. Monte Carlo events generators allow a variation of many kinematic quantities such as cuts on x and/or Q^2 or use of parton showers by the parton shower method introduced in ch. 4. Tuning the used Monte Carlo event generators was a crucial part of this work in order to get a good agreement between data and prediction.

The settings of the Monte Carlo event generator parameters used are given in table 7.1 (RAPGAP) and 7.2 (PYTHIA) (parameters not listed in the tables have then their default values given in [17] and [22]).

| parameter | value | meaning |
|-----------|--------------|--------------------------------|
| XMIN | 0.00001 | x_{min} |
| XMAX | 1.0 | x_{max} |
| QMIN | 1.0 | Q_{min}^2 |
| QMAX | 200.0 | Q_{max}^2 |
| INTE | 0 | electromagnetic interaction |
| IPRO | 12 | process: $eq \rightarrow e'q'$ |
| NFQC | 4 | number of flavours in QCD |
| NFRA | 0 | fragmentation off |
| IFPS | 0, 1, 2 or 3 | switch for parton shower |

Table 7.1: Parameters for RAPGAP

| parameter | value | meaning |
|-----------|--------|--|
| MSEL | 1 | generated process |
| MSTP(91) | 0 or 1 | switch for primordial k_{\perp} |
| MSTP(21) | 2 | γ neutral current only |
| MSTP(11) | 0 | no electron pdf |
| CKIN(23) | 0.0 | x_{min} |
| CKIN(24) | 1.0 | x_{max} |
| CKIN(35) | 1.0 | Q_{min}^2 |
| CKIN(36) | 200.0 | Q_{max}^2 |
| MSTU(101) | 0 | fixed α_{em} |
| MSTJ(41) | 0 or 2 | type of branching allowed in shower |
| MSTP(81) | 0 | multiple interaction off |
| PARP(82) | 0 | regularization scale $p_{\perp 0}$ of the transverse-momentum spectrum for multiple interactions |
| MSTP(61) | 0 or 1 | switch for initial state parton shower |
| MSTP(71) | 0 or 1 | switch for final state parton shower |
| MSTJ(1) | 0 | fragmentation off |
| PARP(67) | 1 | multiplied with Q^2 , to define max. scale for Q^2 -ordered parton shower |
| PARP(71) | 1 | multiplied with Q^2 , to define max. virtuality allowed in time-like showers |

Table 7.2: Parameters for PYTHIA

Chapter 8

Determination of fitting parameters

In this work the PDFs were interfaced using the Les Houches Accord PDF (LHAPDF) interface ([14]). It is a library of PDF sets that can be interfaced with a Monte Carlo event generator.

The PDFs obtained are compared to the PDF set CTEQ6L. This set is a leading order (LO) PDF set, which means that the matrix element of the hard process includes only the first significant process. This is the process $\gamma^*q \rightarrow q$.

Later in this thesis the next to leading order (NLO) set CTEQ6.1M is used for comparison. There the matrix element contains in addition to the LO matrix element $\mathcal{O}(\alpha_s)$ corrections, coming from radiation of partons.

The Monte Carlo event generators used in this work are, like CTEQ6L, of leading order. Hence CTEQ6L is well suited for a comparison with the PDFs obtained by fits using the Monte Carlo generators.

8.1 The CTEQ6L PDF set

The set includes the following PDFs:

- the up-valence quark density

$$xu_v(x, Q^2)$$

- the down-valence quark density

$$xd_v(x, Q^2)$$

- the gluon density

$$xg(x, Q^2)$$

- the ratio

$$x \frac{\bar{u}_s(x, Q^2)}{\bar{d}_s(x, Q^2)}$$

- the sum

$$x(\bar{u}_s(x, Q^2) + \bar{d}_s(x, Q^2))$$

Only the ratio $x \frac{\bar{u}_s(x, Q^2)}{\bar{d}_s(x, Q^2)}$ is parametrized in the functional form:

$$A_f x^{B_f} (1-x)^{C_f} + (1 + D_f x)(1-x)^{E_f}$$

All others are parametrized in the functional form:

$$xf(x, Q_0^2) = A_f x^{B_f} (1-x)^{C_f} e^{x D_f} (1 + e^{E_f} x)^{F_f} \quad (8.1)$$

8.2 Selection of the fitting parameters

Fitting procedures take a lot of computing time which increases with the number of fitting parameters. Since computing time is a major limitation, keeping the number of fitting parameters as small as possible is desirable.

This is why the sensitivity of the χ^2 to the single parameters was investigated before doing the fit by using so called *scans*.

In a scan only one parameter, the scan parameter, is varied within an interval given by the user. The parameter is thereby changed in steps of fix size (given by the user) from the minimum value to the maximum value of the interval. After each step the χ^2 is calculated. All other parameters remain fixed.

If χ^2 has a minimum in the scan, the corresponding value of the scan parameter gives a relatively quick estimate of the value that the parameter will possibly have after doing the fit. How sensitive χ^2 is to the scan parameter (meaning how much influence the scanned parameter has on χ^2) can be seen in how much χ^2 varies when the scan parameter changes.

The goal was to exclude parameters to which χ^2 shows little or even no sensitivity, in order to reduce the number of fitting parameters as much as possible, and by that, to reduce the runtime.

Accordingly, scans for 17 parameters out of the 29 that are contained in the

CTEQ6L set were carried out. The 12 parameters for the valence quark PDFs remained fixed on their starting values during the fit. The reason for this is that at low x the valence quarks play only a minor role and their PDFs have only little influence at low x . In this regime the gluons dominate. Therefore the gluon PDFs and the sea quark PDFs play a significant role here.

The results of the 17 scans will now be discussed.

Fig. 8.1 shows high sensitivity of χ^2 on the gluon parameters and all six parameters for the gluon density $xg(x, Q^2)$ are included in the fit. Fig. 8.3 implies that at least the parameters $A_{\bar{u}+\bar{d}}$, $B_{\bar{u}+\bar{d}}$ and $F_{\bar{u}+\bar{d}}$ lead to strong variation of χ^2 . Parameters $C_{\bar{u}+\bar{d}}$, $D_{\bar{u}+\bar{d}}$ and $E_{\bar{u}+\bar{d}}$ play a less significant role. Nevertheless all six parameters for the sum $x(\bar{u}_S(x, Q^2) + \bar{d}_S(x, Q^2))$ are included in the fit to have the sum completely parametrized. Fig. 8.2 on the other hand shows a very small variation of χ^2 if any at all. This is why all parameters for the ratio $x\frac{\bar{u}_S(x, Q^2)}{\bar{d}_S(x, Q^2)}$ remain fixed at their starting values in the fit. All in all, 12 parameters are included in the fit. This corresponds to $ndf = 92$ degrees of freedom (see ch. 5).

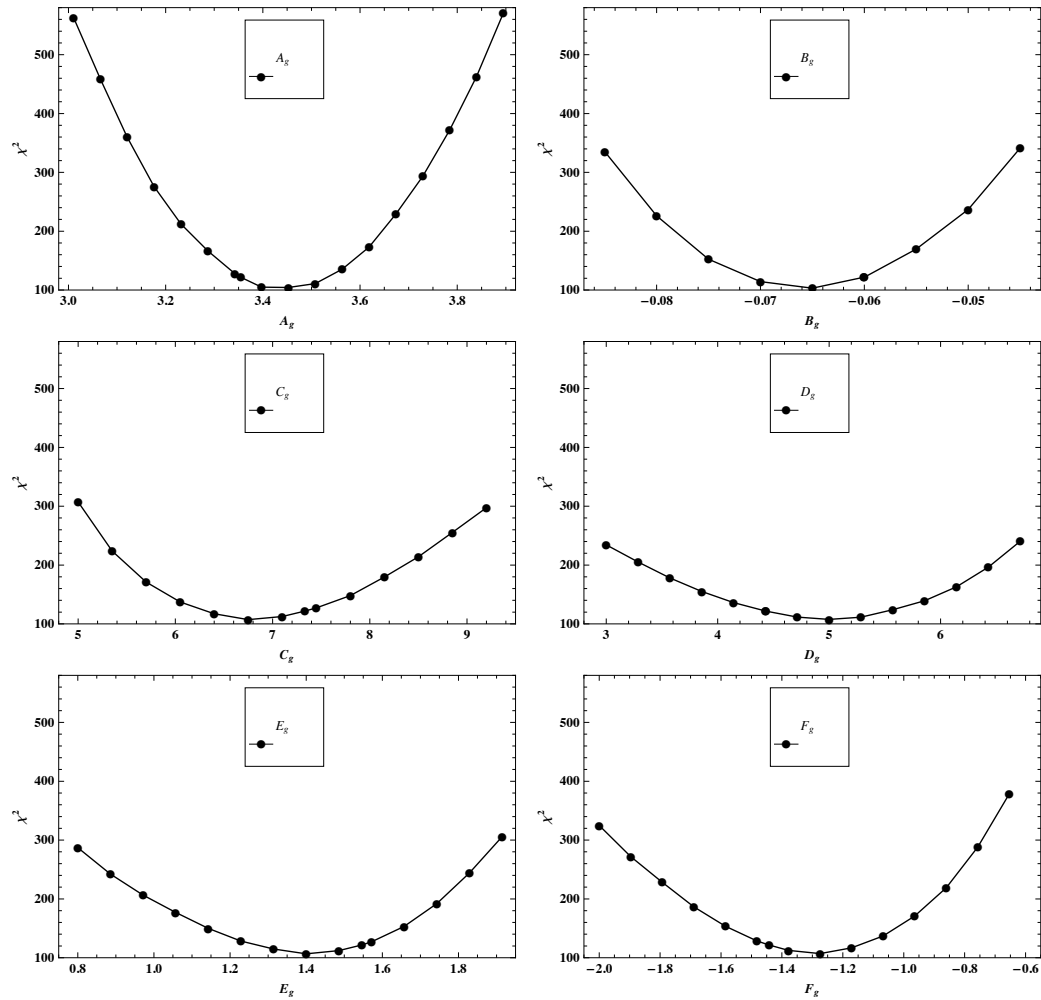


Figure 8.1: Scans of the parameters for the gluon PDF $xg(x, Q^2)$. χ^2 is plotted as a function of the respective parameter value.

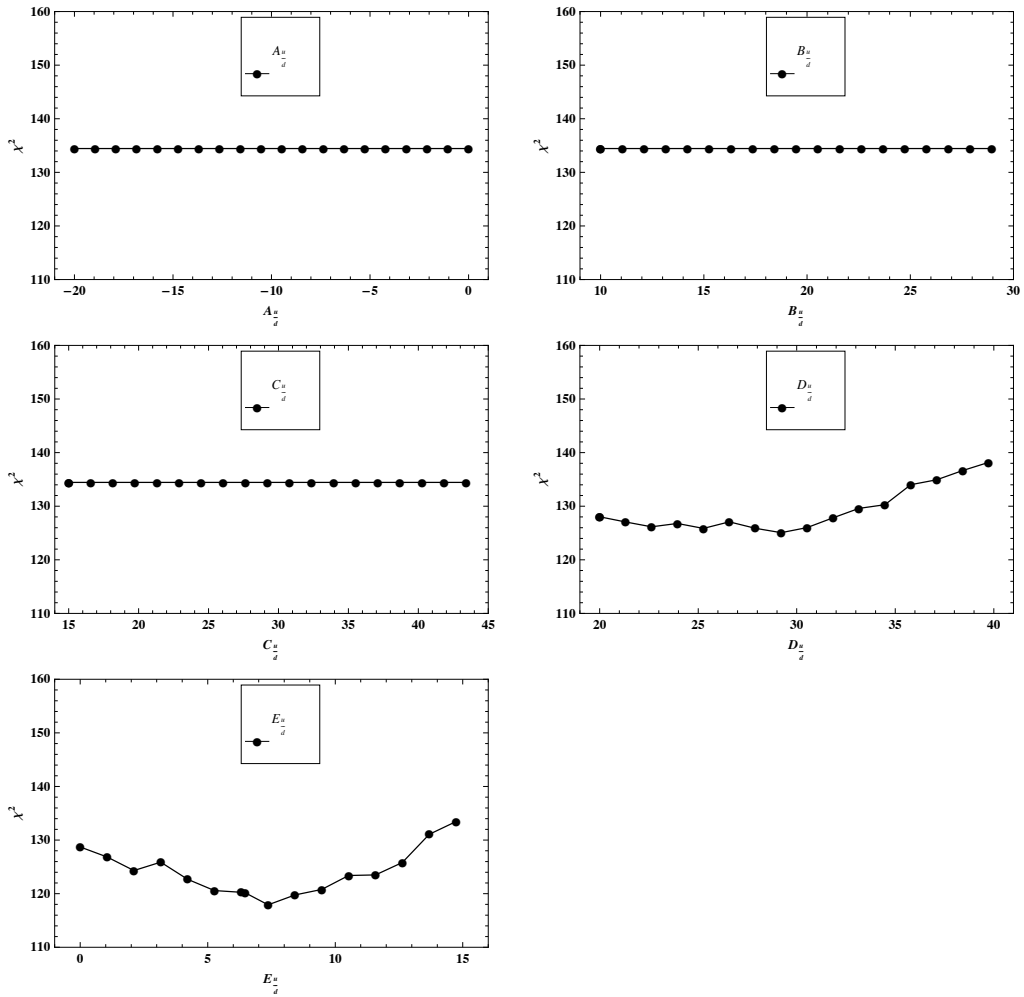


Figure 8.2: Scans of the parameters for the ratio $x \frac{\bar{u}_S(x, Q^2)}{d_S(x, Q^2)}$. χ^2 is plotted as a function of the respective parameter value.

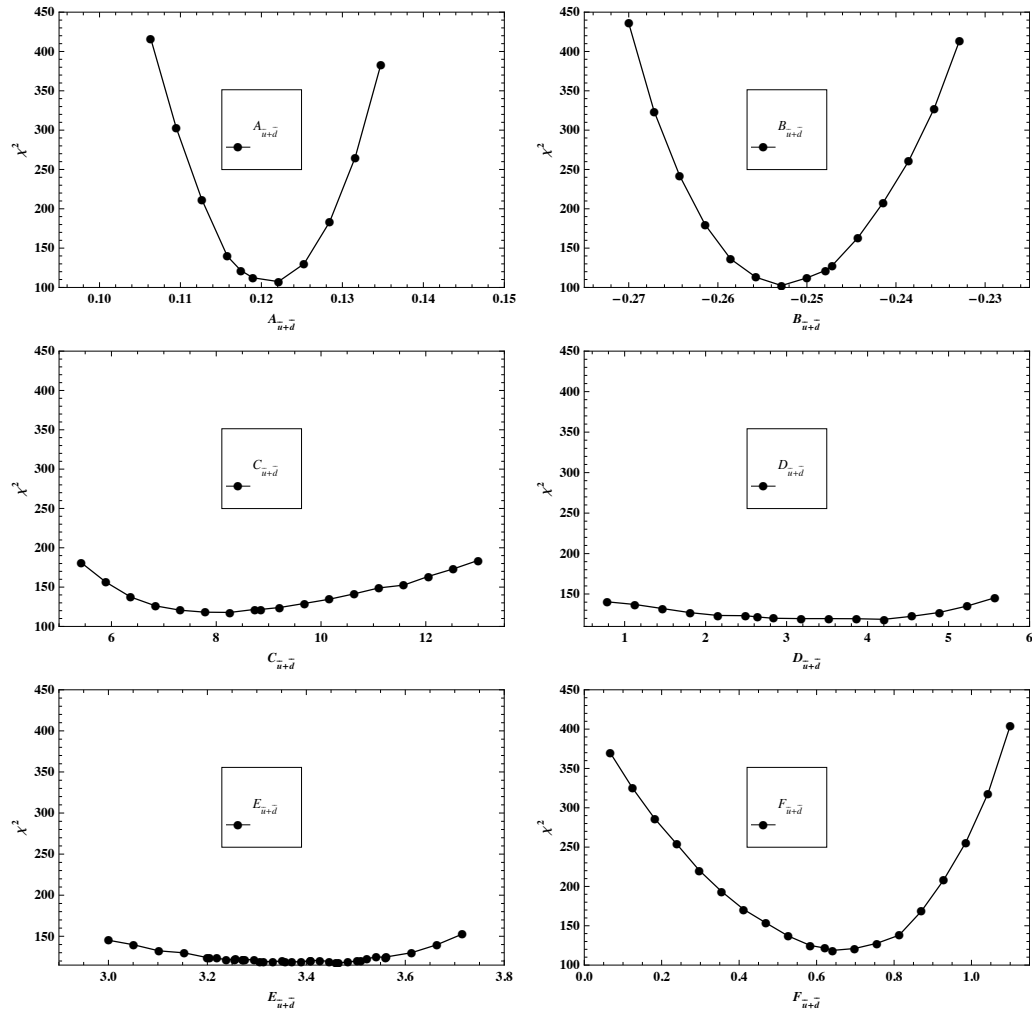


Figure 8.3: Scans of the parameters for the sum $x(\bar{u}_S(x, Q^2) + \bar{d}_S(x, Q^2))$. χ^2 is plotted as a function of the respective parameter value.

The predictions of the Monte Carlo event generators using the settings of table 7.2 and 7.1 are shown in figs 8.4 and 8.5.

In fig. 8.4 the structure function $F_2(x, Q^2)$ as a function of x for different values of Q^2 is shown. The agreement between data and the Monte Carlo event generators is very good with $\chi^2/ndf = 0.84$ for RAPGAP and $\chi^2/ndf = 0.96$ for PYTHIA.

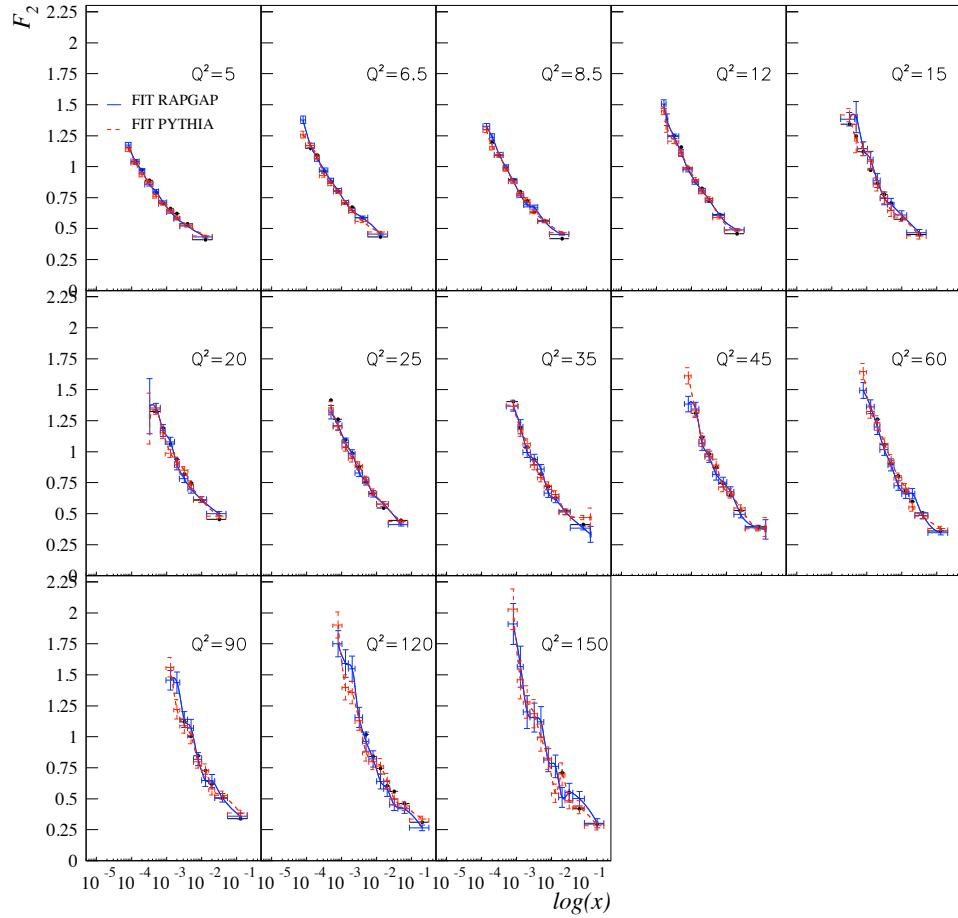


Figure 8.4: The structure function $F_2(x, Q^2)$ as a function of x for values from $Q^2 = 5 \text{ GeV}^2$ to $Q^2 = 150 \text{ GeV}^2$. The data is given as dots and the predictions coming from RAPGAP and PYTHIA are given by points connected by a full line and a broken line respectively.

In fig. 8.5 the structure function $F_2(Q^2, x)$ is plotted as a function of Q^2 for different values of x for (a) RAPGAP and (b) PYTHIA. In the double logarithmic plot the data are represented by dots and the predictions by a full line. The slight waves in the lines come from the predictions being of statistical nature.

One can see here the predicted scaling violation (see sec. 2.3), namely that $F_2(Q^2, x)$ is a function of x and Q^2 .

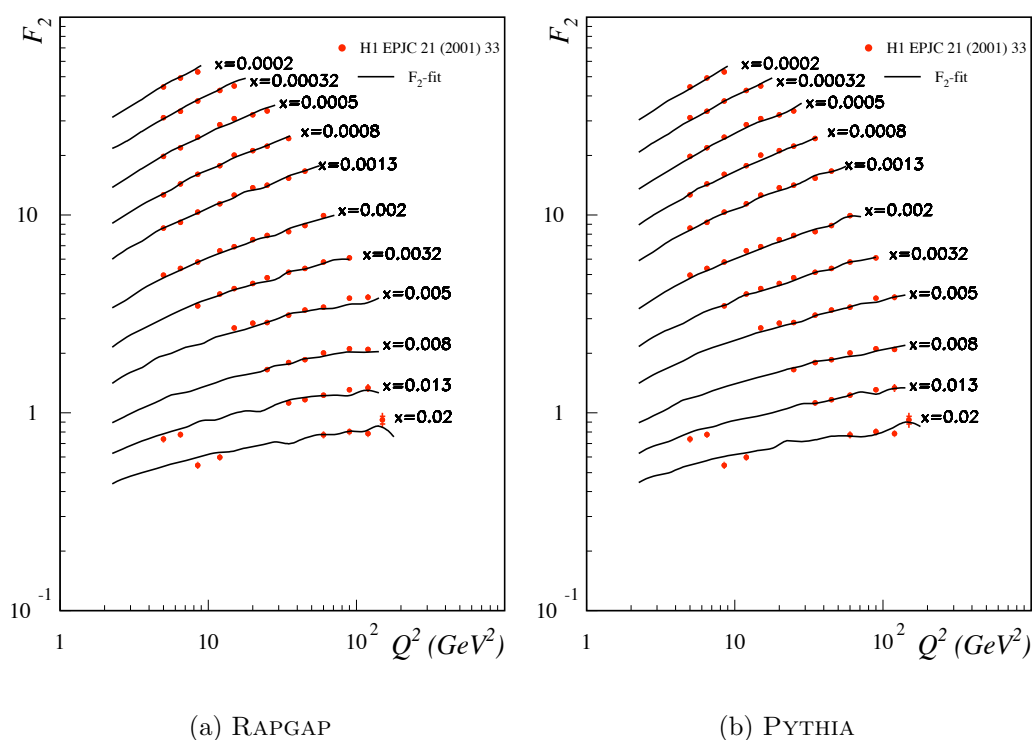
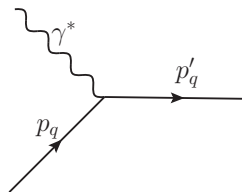


Figure 8.5: F_2 prediction plotted as a function of Q^2 and from $x = 0.0002$ to $x = 0.02$ for (a) RAPGAP and (b) PYTHIA. The dots represent the data, the full line the prediction

Chapter 9

The influence of parton showers on PDFs

In the main process



4-momentum conservation gives (neglecting parton masses)

$$\begin{aligned}(q + p_q)^2 &= p_q'^2 \\ -Q^2 + 2qp_q &= 0\end{aligned}\tag{9.1}$$

where p_q is the quark 4-momentum, which carries a fraction ξ of the protons momentum p ($p_q = \xi p$). This gives then:

$$\begin{aligned}-Q^2 + 2q \cdot \xi p &= 0 \\ \Rightarrow \xi &= \frac{Q^2}{2qp} = x_{Bj}\end{aligned}\tag{9.2}$$

which is eq. (2.13).

Using the Parton Shower Method described in ch. 4 in a Monte Carlo event generator has the effect that the transversal momentum k_\perp and the virtuality of the parton, arising in one branching are no longer neglected in the next branching. Therefore, virtualities of the parton have now to be taken into

account. The initial state parton showers lead to spacelike virtualities < 0 and final state parton showers to timelike virtualities > 0 . But then

$$p_q^2 = -k^2 (< 0) \quad (9.3)$$

$$p_q'^2 = m^2 (> 0) \quad (9.4)$$

Equivalently to the derivation of eq. (9.2) one gets:

$$\begin{aligned} (q + p_q)^2 &= p_q'^2 \\ -Q^2 - k^2 + 2qp_q &= m^2 \\ \Rightarrow \xi &= \frac{Q^2 + k^2 + m^2}{2qp} \neq x_{Bj} \end{aligned} \quad (9.5)$$

where $p_q = \xi p$ was used again.

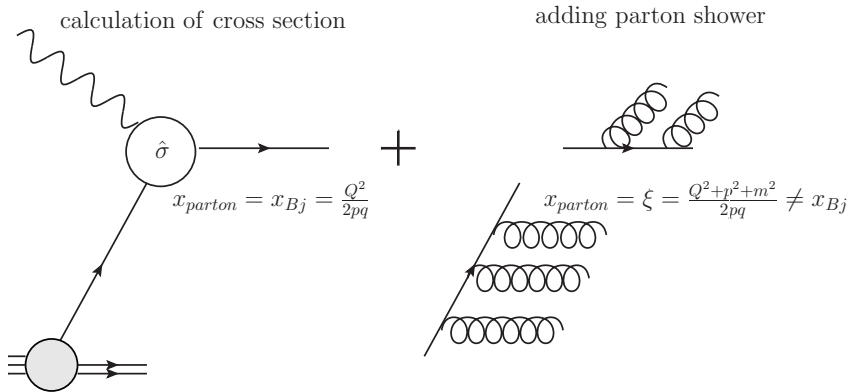


Figure 9.1: First the cross section is calculated and then parton shower added

To simulate a process, a Monte Carlo event generator calculates first the total cross section σ_{tot} . The differential cross section is then calculated by:

$$\frac{d\sigma}{dQ^2 dx} = \frac{N_{ev}(x, Q^2)}{N_{tot}} \cdot \sigma_{tot}. \quad (9.6)$$

$N_{ev}(x, Q^2)$ is the number of events in a distinctive x - and Q^2 -bin. The Monte Carlo event generator then adds parton shower in accordance to the Parton Shower Method described in sec. 4.2. This changes the momentum fraction of the parton x_p given in eq. (9.2) to the one given in eq. 9.5 (see fig. 9.1). As a consequence, events that are in a certain x -bin migrate to other x -bins, when parton shower are added. This means, that the $N_{ev}(x, Q^2)$ for the

single x -bins change, which leads to a change of $F_2(x, Q^2) \propto Q^4 x \frac{d^2\sigma}{dQ^2 dx}$. If the Monte Carlo event generator, running without parton shower, describes $F_2(x, Q^2)$ properly according to eq. (3.3), then

$$F_2(x, Q^2) \propto \left(\frac{d^2\sigma}{dQ^2 dx} \right)^{no\ ps}. \quad (9.7)$$

In this case, the description of the Monte Carlo event generator, running with parton shower, must be corrected to:

$$F_2(x, Q^2) \propto \left(\frac{d^2\sigma}{dQ^2 dx} \right)^{ps} \cdot \frac{\left(\frac{d^2\sigma}{dQ^2 dx} \right)^{no\ ps}}{\left(\frac{d^2\sigma}{dQ^2 dx} \right)^{ps}}. \quad (9.8)$$

The ratio on the left side of the last equation corrects for the migration effects on x due to parton shower.

Since the x -dependent PDF enters the cross section, the correction can be done by changing the PDF. To do so, one can try to obtain a PDF by a fit of F_2 -data to predictions for $F_2(x, Q^2)$, coming from a Monte Carlo event generator, running with parton shower.

Monte Carlo event generators running with parton shower and employing a PDF obtained without parton shower (like CTEQ6L), will not describe data properly. Since parton shower are needed to describe exclusive measurements, a correction e.g. in form of an optimized PDF set is desirable.

In all this the Monte Carlo event generators RAPGAP and LEPTO make an exception. Here the splitting variables for the parton shower are selected with the constraint, that $x_p = x_{Bj}$ for a given Q^2 . The inclusion of parton shower will therefore not lead to a change of $F_2(Q^2, x)$.

The influence of parton shower on the PDFs, obtained by fits of Monte Carlo event generator predictions to F_2 -data, is investigated in this thesis. This is done on the basis of three different scenarios:

1. RAPGAP and PYTHIA without parton shower

PDFs are fitted using RAPGAP and PYTHIA, both with parton showers switched off. In RAPGAP this is realized by the switch IFPS=0 and in PYTHIA by the switches MSTP(61)=0, MSTP(71)=0 and MSTP(91)=0 (see tables 7.1 and 7.2).

These fits can then be compared to the LO PDF set CTEQ6L and with each other. If it is at all possible to fit PDFs using Monte Carlo generators, all three PDFs should coincide, because without parton showers,

the kinematics is left unchanged and the two generators should then yield the same result as CTEQ6L.

2. RAPGAP with parton shower

PDFs are fitted using RAPGAP. Parton showers are included in the generation process. This happens step by step by first including only initial state parton shower, then only final state parton shower and finally the combination of initial state and final state parton showers. In RAPGAP this is realized by the switch IFPS=1/2/3. By using a method as described in [3], RAPGAP does not change x , when parton shower are included. Hence there should be no difference between parton shower being switched on or off.

3. PYTHIA with parton shower

PDFs are fitted using PYTHIA. Parton showers are included in the generation process and switched on stepwise. First only an intrinsic k_{\perp} is included¹. Then initial state parton shower, final state parton shower and finally the combination of initial and final state parton shower are switched on. In PYTHIA this is realized by the switches MSTP(61)=0/1, MSTP(72)=0/1 and MSTP(91)=0/1.

Unlike RAPGAP, PYTHIA changes x , when parton shower are included. This means that in this case a change in the resulting PDFs is expected.

After that a comparison of the PDFs resulting from the fits with the PDF set CTEQ6.1 M is made. Finally the χ^2 values for RAPGAP's and PYTHIA's data description when using CTEQ6L are compared to the corresponding χ^2 values obtained when using the fitted PDFs. For the sake of readability the following terms are defined:

- nops-PDF: PDF obtained by a fit of Monte Carlo event generator-predictions to F_2 -data using the Monte Carlo event generator without parton shower
- ikt-PDF: PDF obtained by a fit of Monte Carlo event generator-predictions to F_2 -data using the Monte Carlo event generator with intrinsic k_{\perp}
- ips-PDF: PDF obtained by a fit of Monte Carlo event generator-predictions to F_2 -data using the Monte Carlo event generator with initial state parton shower

¹Intrinsic (or primordial) k_{\perp} expresses the idea, that the partons contained in the proton can, due to fermi motion, have a certain transversal momentum k_{\perp} , while the proton as such only carries longitudinal momentum (i.e. in beam direction). This k_{\perp} is modeled as a Gaussian in k_{\perp}^2 , having a width of $2 \text{ GeV}/c$.

- fps-PDF: PDF obtained by a fit of Monte Carlo event generator-predictions to F_2 -data using the Monte Carlo event generator with final state parton shower
- ifps-PDF: PDF obtained by a fit of Monte Carlo event generator-predictions to F_2 -data using the Monte Carlo event generator with the combination of initial and final state parton shower

9.1 RAPGAP and PYTHIA excluding parton shower

Fig. 9.2 shows the fitted PDF of the gluons, fig. 9.3 the up sea quarks² and fig. 9.4 the down sea quarks³ as a function of x obtained by a fit using RAPGAP and the PDF set CTEQ6L for $Q^2 = 10, 100, 1000$ and 10000 GeV².

Fig. 9.5 shows the fitted PDF of the gluons, fig. 9.6 the up sea quarks and fig. 9.7 the down sea quarks as a function of x obtained by a fit using PYTHIA and the PDF set CTEQ6L.

In both cases the curve representing CTEQ6L and that representing the fitted PDFs coincide, such that they are barely distinguishable in the plot.

Both generators show very good agreement with CTEQ6L (the curves are barely distinguishable) and with each other over four orders of magnitude of Q^2 .

The parameter values resulting from the fit are given in the tables B.1a and B.3a.

Interpretation

The observed effects can be understood as follows. If parton showers are absent, both Monte Carlo event generators simply use the evolved PDF at the hard scale Q^2 . Hence the fits of the Monte Carlo generator, which use LO matrix elements for the hard process, should coincide with the PDF set CTEQ6L, that is obtained using the evolved PDF and LO matrix elements as well. This is verified by the agreement between the PDFs resulting from the fits and the CTEQ6L set.

²Up sea quark PDFs stands here for the sum of up sea quark and charm sea quark PDFs

³Down sea quark PDFs stands here for the sum of down sea quark, strange sea quark and bottom sea quark PDFs

This shows, that PDF fits using Monte Carlo event generators lead to sensible and consistent results and that the proposed method to fit PDFs using Monte Carlo generators can be applied successfully.

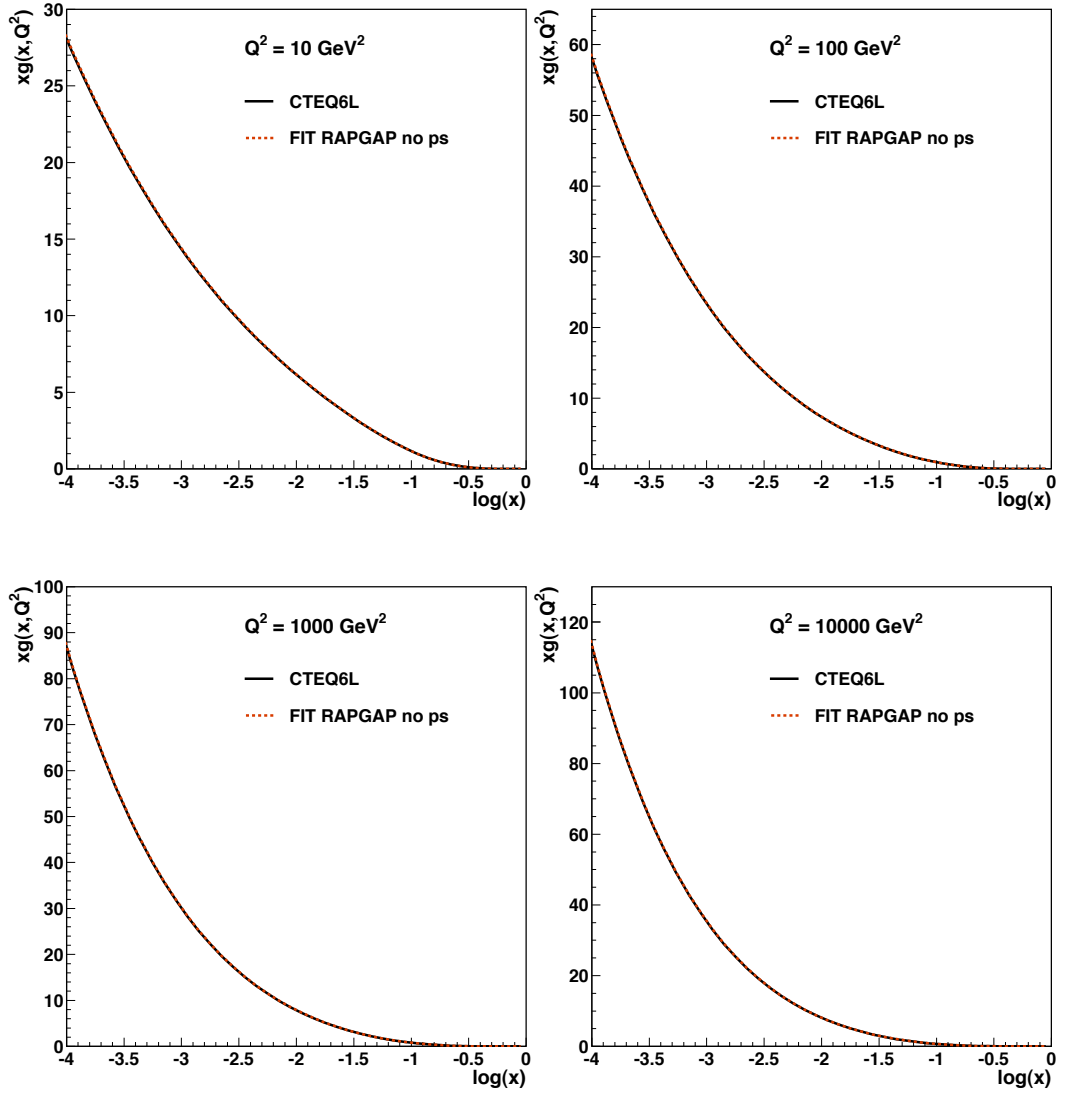


Figure 9.2: Gluon PDF as a function of x from CTEQ6L (black, full line) and from RAPGAP (red, broken line), fitted with no parton showers, at $Q^2 = 10 \text{ GeV}^2$, 100 GeV^2 , 1000 GeV^2 and 10000 GeV^2 .

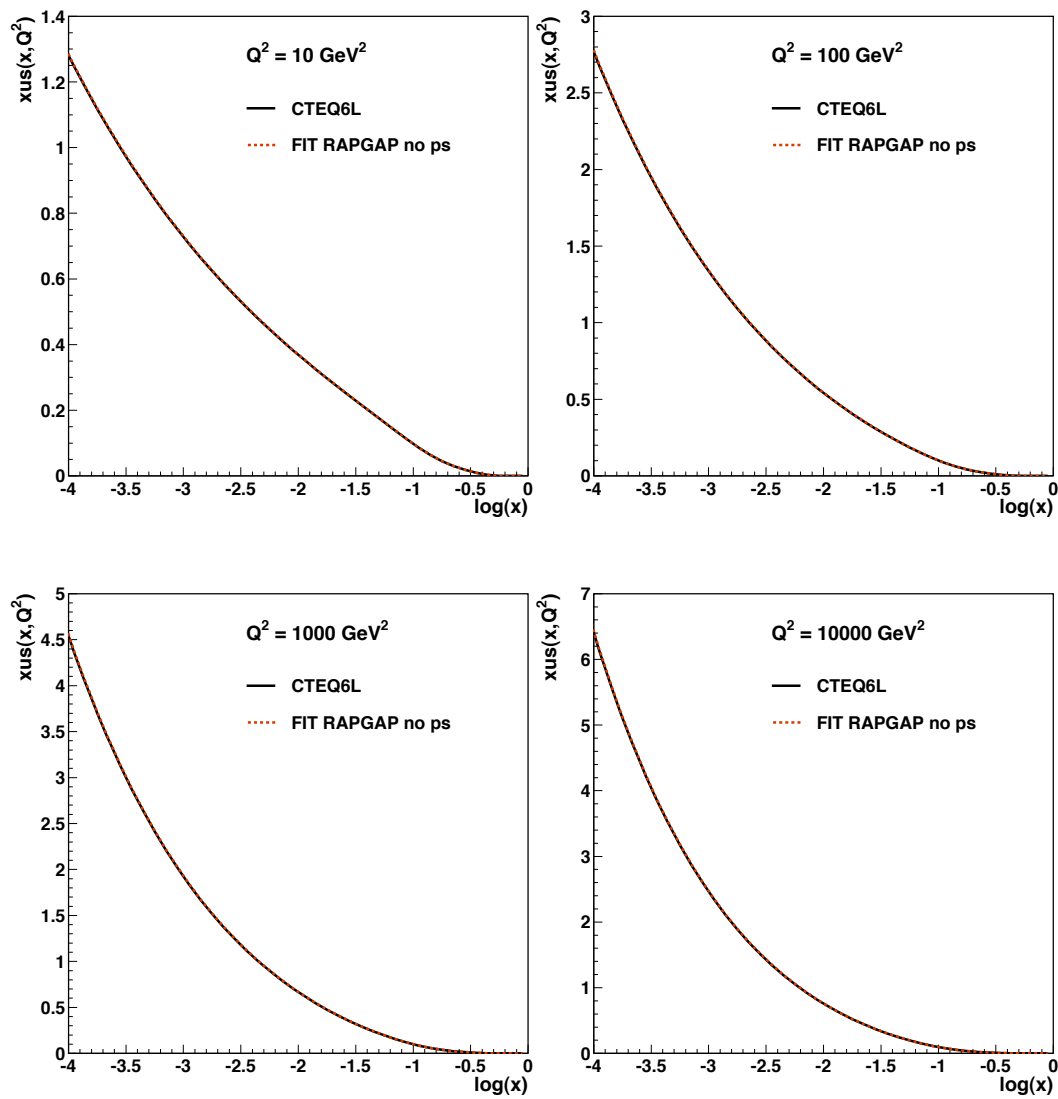


Figure 9.3: Up sea quark PDF as a function of x from CTEQ6L (black, full line) and from RAPGAP (red, broken line), fitted with no parton showers, at $Q^2 = 10 \text{ GeV}^2$, 100 GeV^2 , 1000 GeV^2 and 10000 GeV^2 .

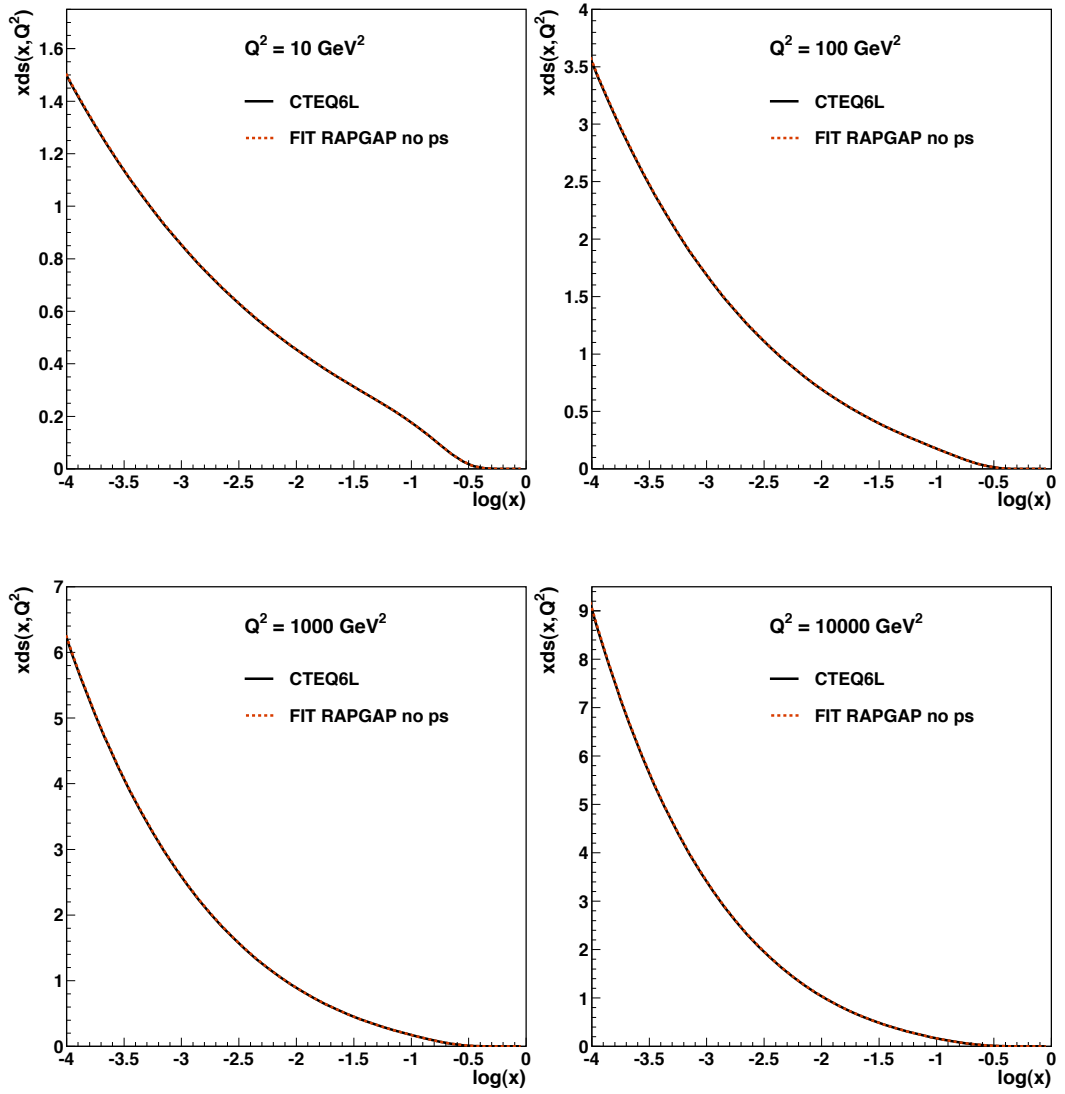


Figure 9.4: Down sea quark PDF as a function of x from CTEQ6L (black, full line) and from RAPGAP (red, broken line), fitted with no parton showers, at $Q^2 = 10 \text{ GeV}^2$, 100 GeV^2 , 1000 GeV^2 and 10000 GeV^2 .

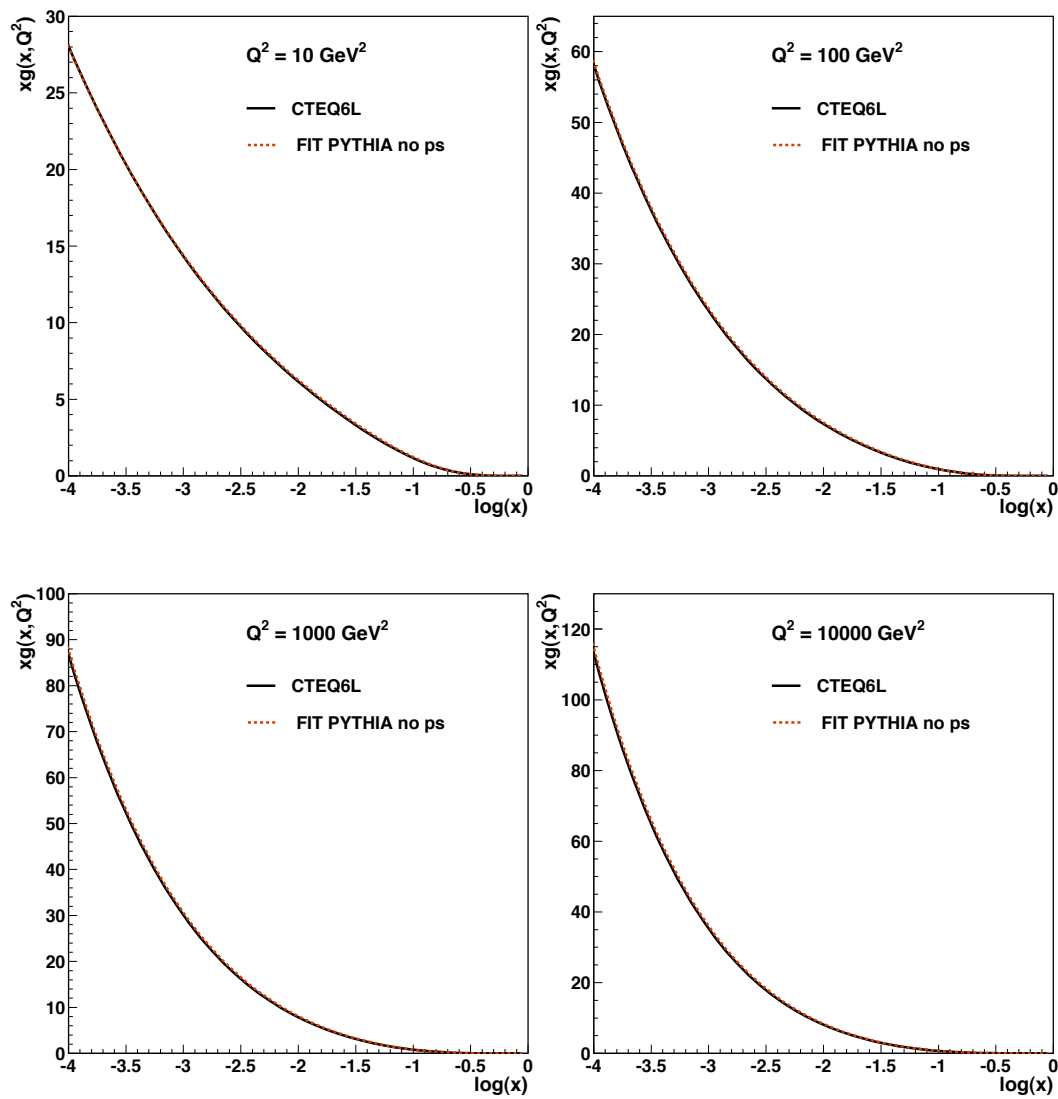


Figure 9.5: Gluon PDF as a function of x from CTEQ6L (black, full line) and from PYTHIA (red, broken line), fitted with no parton showers, at $Q^2 = 10 \text{ GeV}^2$, 100 GeV^2 , 1000 GeV^2 and 10000 GeV^2 .

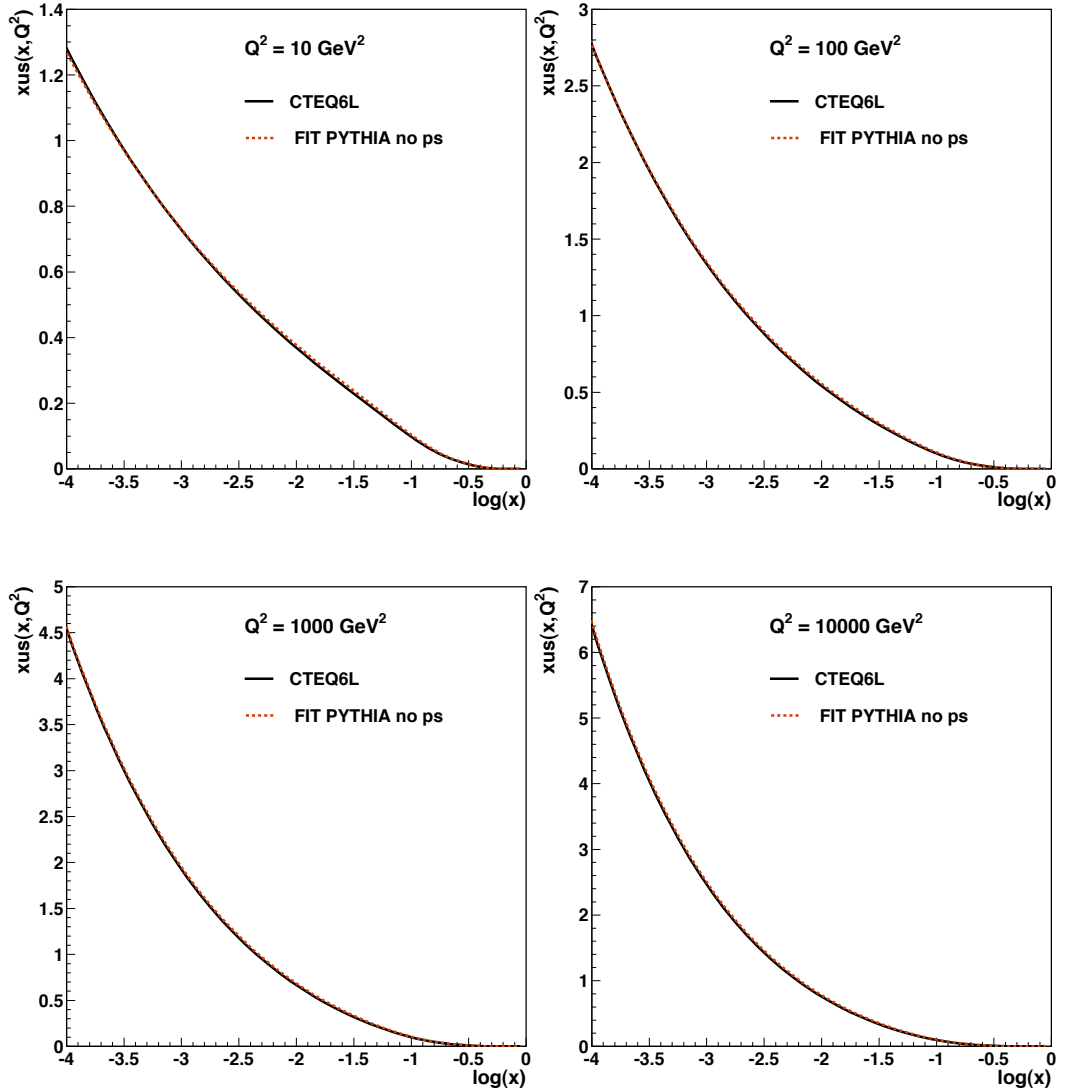


Figure 9.6: Up sea quark PDF as a function of x from CTEQ6L (black, full line) and from PYTHIA (red, broken line), fitted with no parton showers, at $Q^2 = 10 \text{ GeV}^2$, 100 GeV^2 , 1000 GeV^2 and 10000 GeV^2 .

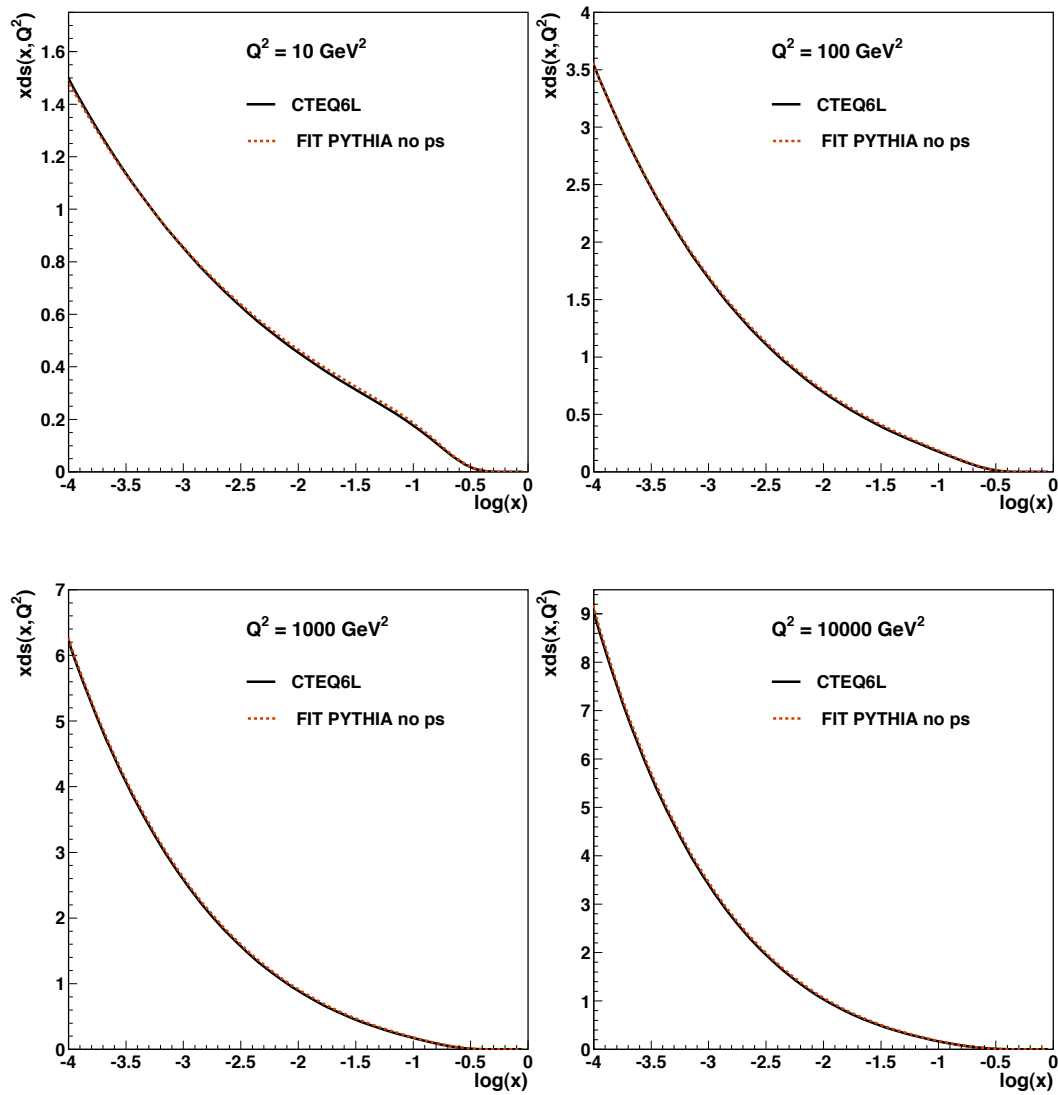


Figure 9.7: Down sea quark PDF as a function of x from CTEQ6L (black, full line) and from PYTHIA (red, broken line), fitted with no parton showers, at $Q^2 = 10 \text{ GeV}^2$, 100 GeV^2 , 1000 GeV^2 and 10000 GeV^2 .

9.2 RAPGAP including parton showers

Next the influence of parton showers in RAPGAP on the PDFs resulting from the fit are investigated. First initial state parton shower, then final state parton shower and finally the combination of initial and final state parton showers were included in the fit.

Fig. 9.8 shows the gluon, fig. 9.9 the up sea quark and fig. 9.10 the down sea quark PDFs as a function of x resulting from the fit using RAPGAP for $Q^2 = 10, 100, 1000$ and 10000 GeV^2 .

For low Q^2 the ips-PDFs of the up sea quarks and the down sea quarks are slightly below the respective nops-PDFs at low and high x . The SIMPLEX method gives no errors for the fitted parameters of the PDFs. Therefore the significance of these deviations cannot be determined.

The ifps-PDFs coincide with the nops-PDFs for the gluon, the up sea quark and the down sea quark at all values of Q^2 .

The parameter values resulting from the fit are given in the tables B.1.

Interpretation

Since in RAPGAP x is not changed by parton showers, x has the same value as in the case where parton shower are switched off. Hence the same explanation as in sec. 9.1 can be given for the observed fact, that the PDFs resulting from the fits including parton shower coincide with the nops-PDFs (and regarding to sec. 9.1 therefore to CTEQ6L).

It shows that in RAPGAP the kinematic variables x and Q^2 are indeed independent of the inclusion of parton showers in the generation process. This is the first time that this characteristic of RAPGAP is directly proven by a systematic investigation of the effect of parton showers on quantities sensitive to the kinematics as are PDFs.

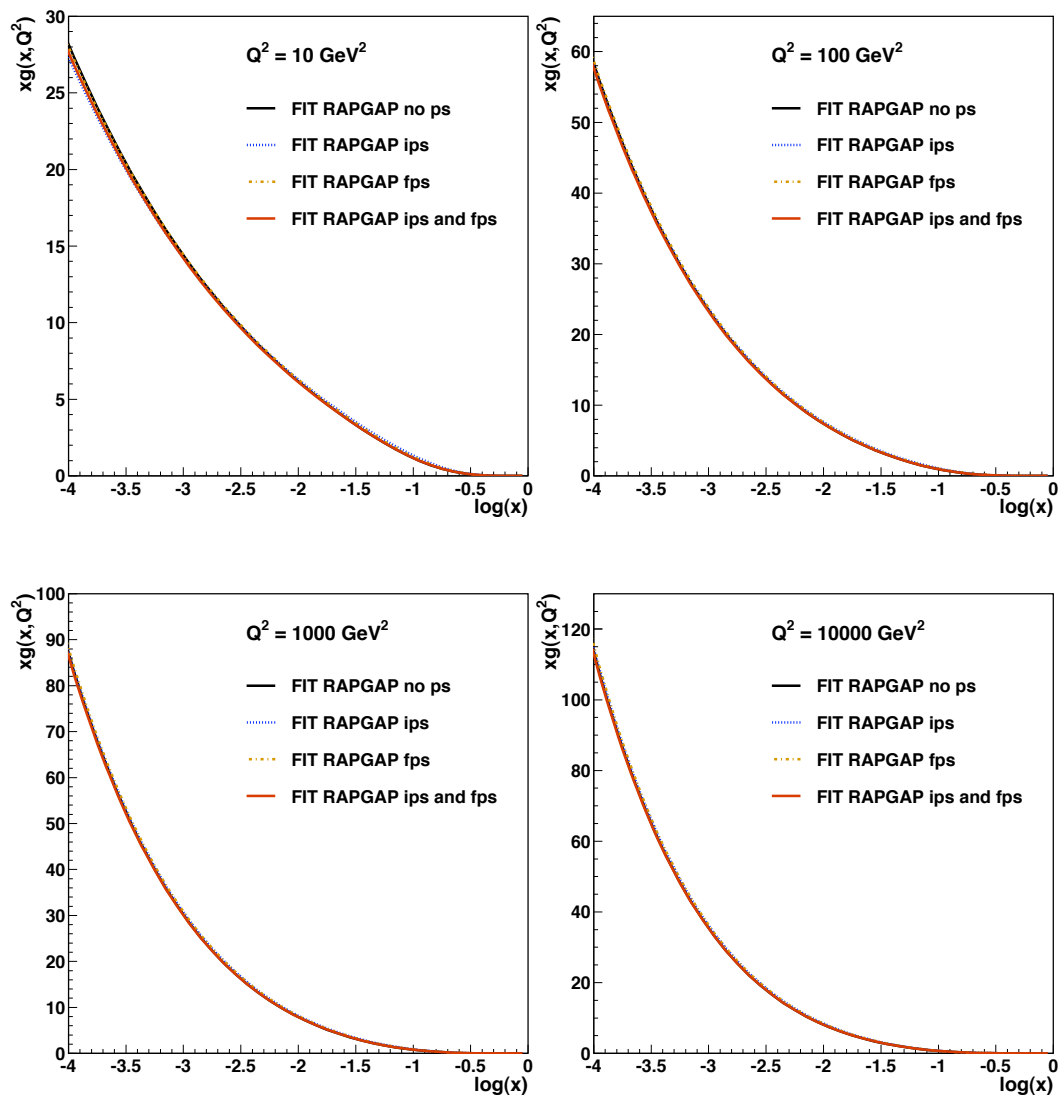


Figure 9.8: Gluon PDF as a function of x from RAPGAP, fitted with no parton shower, initial state parton shower, final state parton shower and the combination of initial and final state parton shower, at $Q^2 = 10 \text{ GeV}^2$, 100 GeV^2 , 1000 GeV^2 and 10000 GeV^2 .

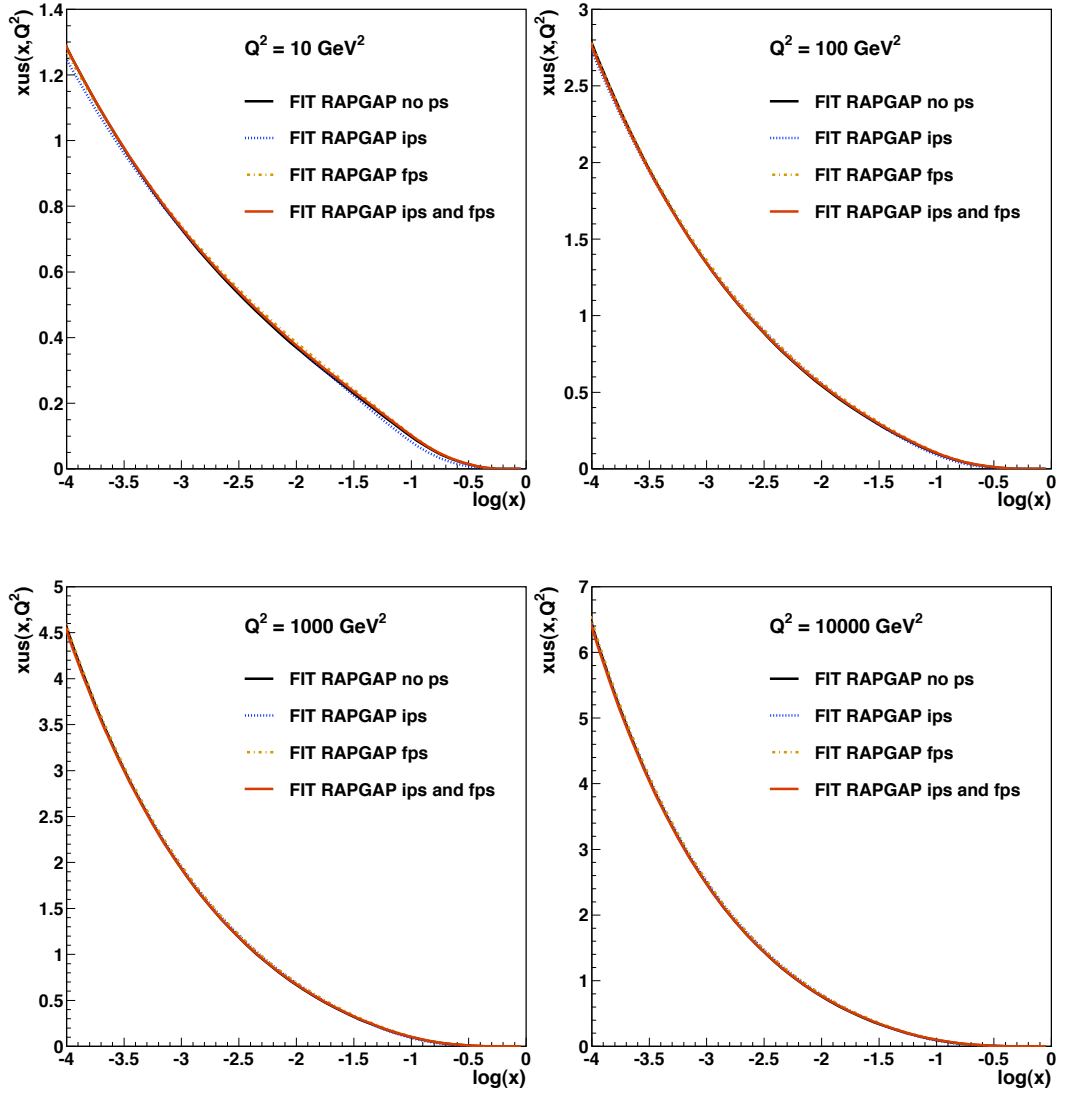


Figure 9.9: Up sea quark PDF as a function of x from RAPGAP, fitted with no parton shower, initial state parton shower, final state parton shower and the combination of initial and final state parton shower, at $Q^2 = 10 \text{ GeV}^2$, 100 GeV^2 , 1000 GeV^2 and 10000 GeV^2 .

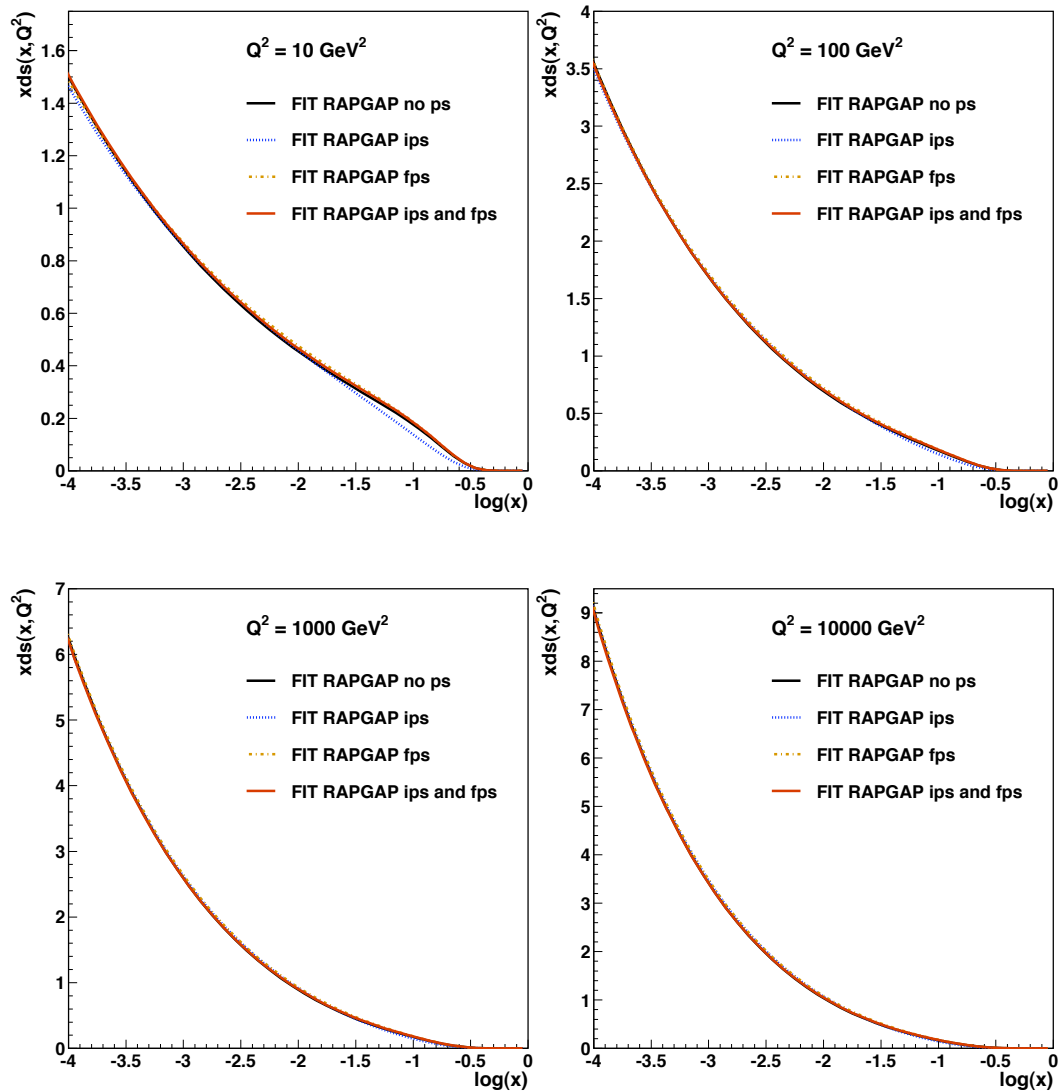


Figure 9.10: Down sea quark PDF as a function of x from RAPGAP, fitted with no parton shower, initial state parton shower, final state parton shower and the combination of initial and final state parton shower, at $Q^2 = 10 \text{ GeV}^2$, 100 GeV^2 , 1000 GeV^2 and 10000 GeV^2 .

9.3 PYTHIA including parton showers

In PYTHIA there are two options to order the parton showers (see ch. 4) - k_{\perp} -ordered or Q^2 -ordered. The effects of both orderings are investigated in the following.

9.3.1 k_{\perp} -ordered parton showers

Fig. 9.11 shows the gluon, fig. 9.12 the up sea quark and fig. 9.13 the down sea quark PDFs as a function of x resulting from the fit using PYTHIA for $Q^2 = 10, 100, 1000$ and 10000 GeV^2 . The parton shower were k_{\perp} -ordered.

First only intrinsic k_{\perp} was taken into account, then in addition initial state parton shower, then final state parton shower and at the end the combination of initial and final state parton shower.

In the x region below 10^{-2} the ik_{\perp} -PDFs, ips-PDFs, fps-PDFs and ifps-PDFs of gluons, up sea quarks and down sea quarks decrease in comparison to the respective nops-PDFs. The order (going from lowest decrease to highest decrease compared to the nops-PDF) is hereby ik_{\perp} -PDFs, fps-PDFs, ips-PDFs and ifps-PDFs. For $x \approx 10^{-4}$ and $Q^2 = 10 \text{ GeV}^2$ the ips-PDF and the fps-PDF of the up sea quark and the down sea quark cross.

In the x region above 10^{-2} the ik_{\perp} -PDFs, ips-PDFs, fps-PDFs and ifps-PDFs of gluons, up sea quarks and down sea quarks increase in comparison to the respective nops-PDFs. The ik_{\perp} -PDFs of the up sea quark and the down sea quark show a slight decrease as compared to the respective nops-PDFs. The ik_{\perp} -PDF of the gluon coincides with the respective nops-PDF.

For increasing Q^2 , all PDFs show a clear reduction of the decrease observed for the inclusion of parton showers and intrinsic k_{\perp} .

The parameter values resulting from the fit are given in the tables B.2.

Interpretation

In PYTHIA, where no reconstruction of the kinematics is applied, parton shower change x according to eq. (9.5) due to the virtualities of the particle obtained by parton shower in the initial and final state.

The more parton showers are allowed and the more virtuality can be gained by each shower, the stronger the change in x will be. The inclusion of an intrinsic k_{\perp} in the event generation, which gives some k_{\perp} to the partons and thereby some virtuality right from the start, leads to a first change when the

Monte Carlo prediction is fitted to F_2 -data. When additionally initial state and final state parton shower are switched on, even more virtuality is gained by the parton shower. Hence x is changed even more as is evident from the observed stepwise decrease of the PDFs at low x .

The decrease of the effects due to parton shower and intrinsic k_\perp for increasing Q^2 becomes clear in eq. (9.5). For sufficiently high Q^2 , $k_\perp^2 \ll Q^2$ and $m^2 \ll Q^2$ can be neglected, which leads back to eq. (9.2).

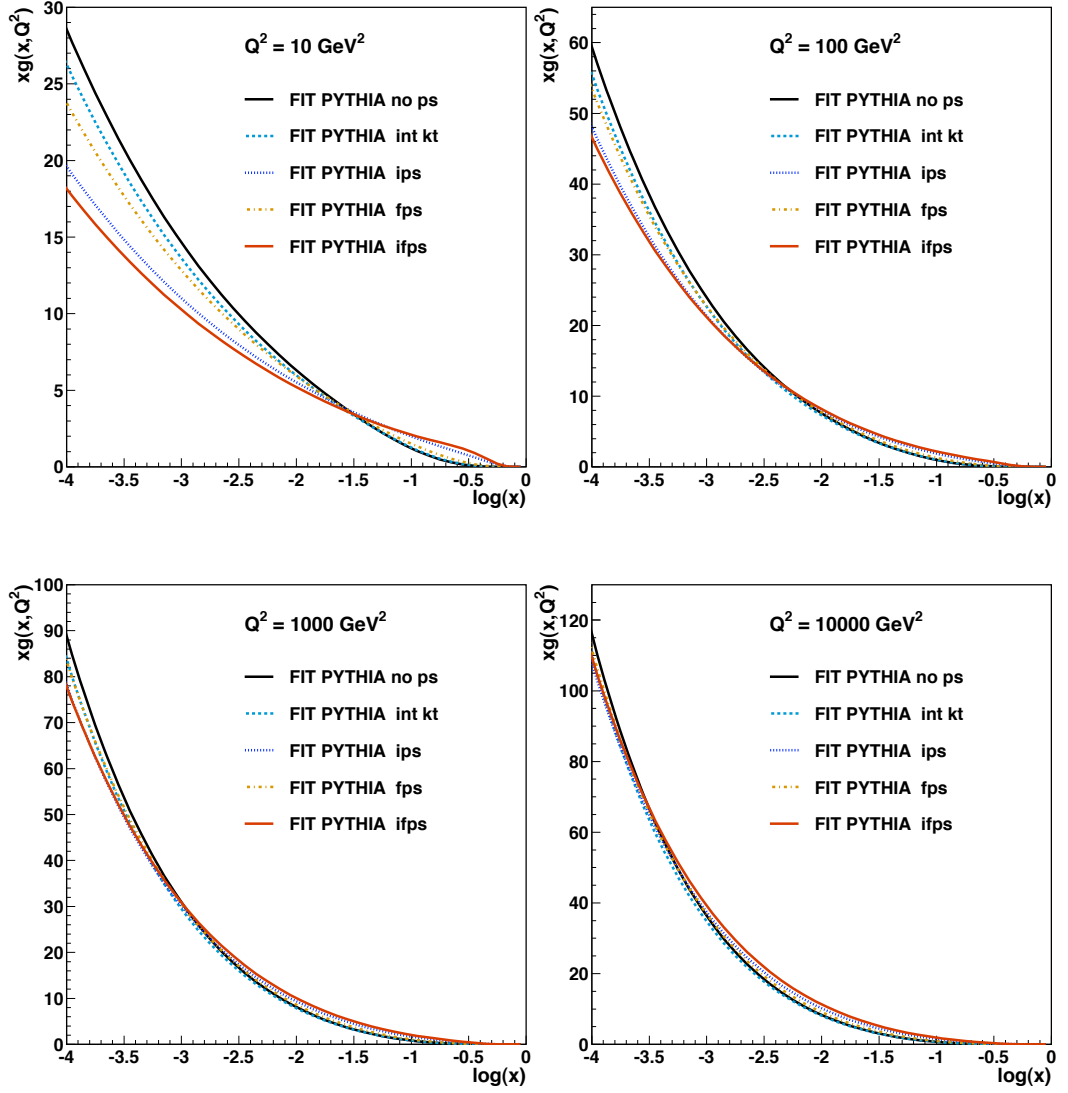


Figure 9.11: Gluon PDF as a function of x from k_{\perp} -ordered PYTHIA, fitted with no parton shower, no parton shower + intrinsic k_{\perp} , initial state parton shower, final state parton shower and the combination of initial and final state parton shower, at $Q^2 = 10 \text{ GeV}^2$, 100 GeV^2 , 1000 GeV^2 and 10000 GeV^2 .

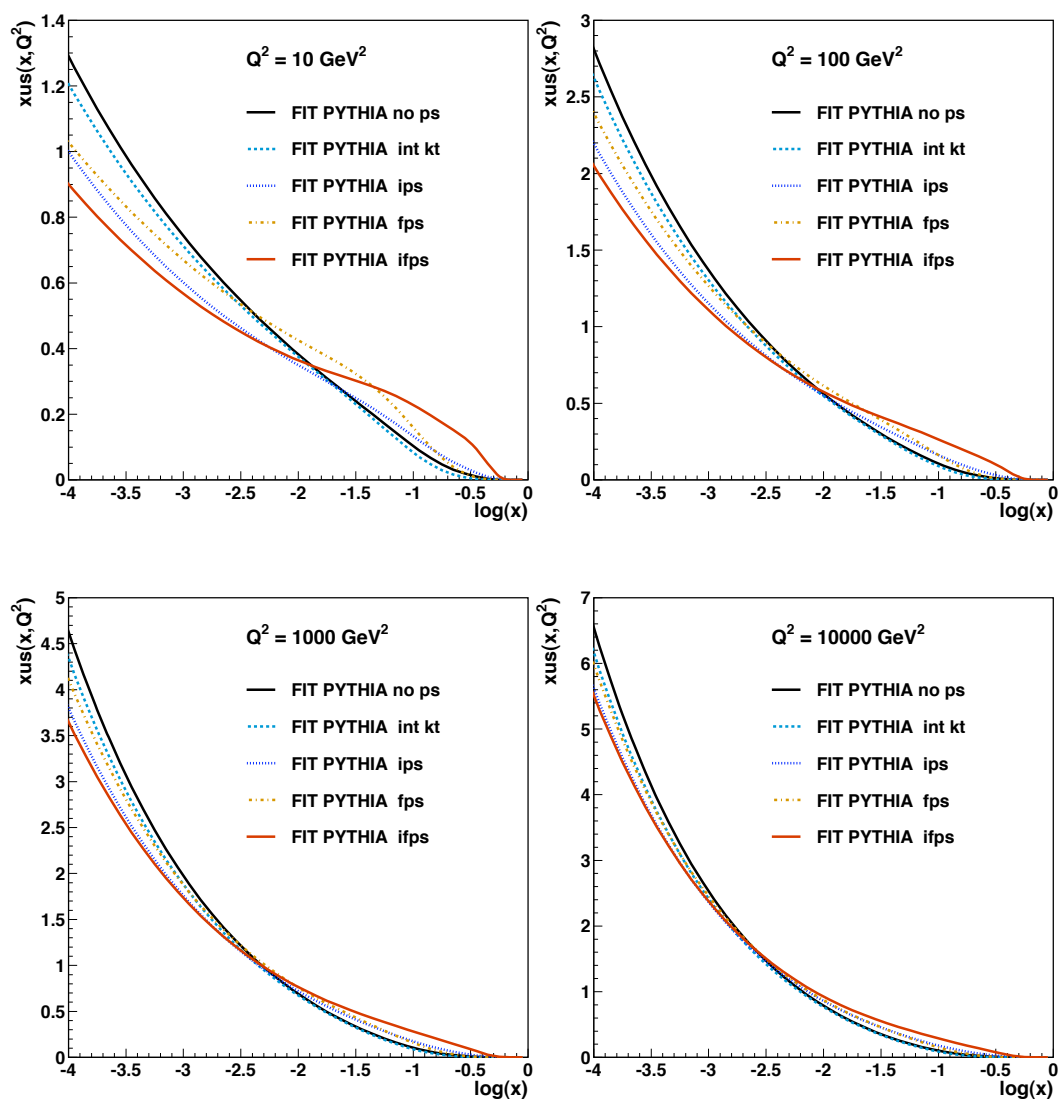


Figure 9.12: Up sea quark PDF as a function of x from k_{\perp} -ordered PYTHIA, fitted with no parton shower, no parton shower + intrinsic k_{\perp} , initial state parton shower, final state parton shower and the combination of initial and final state parton shower, at $Q^2 = 10 \text{ GeV}^2$, 100 GeV^2 , 1000 GeV^2 and 10000 GeV^2 .

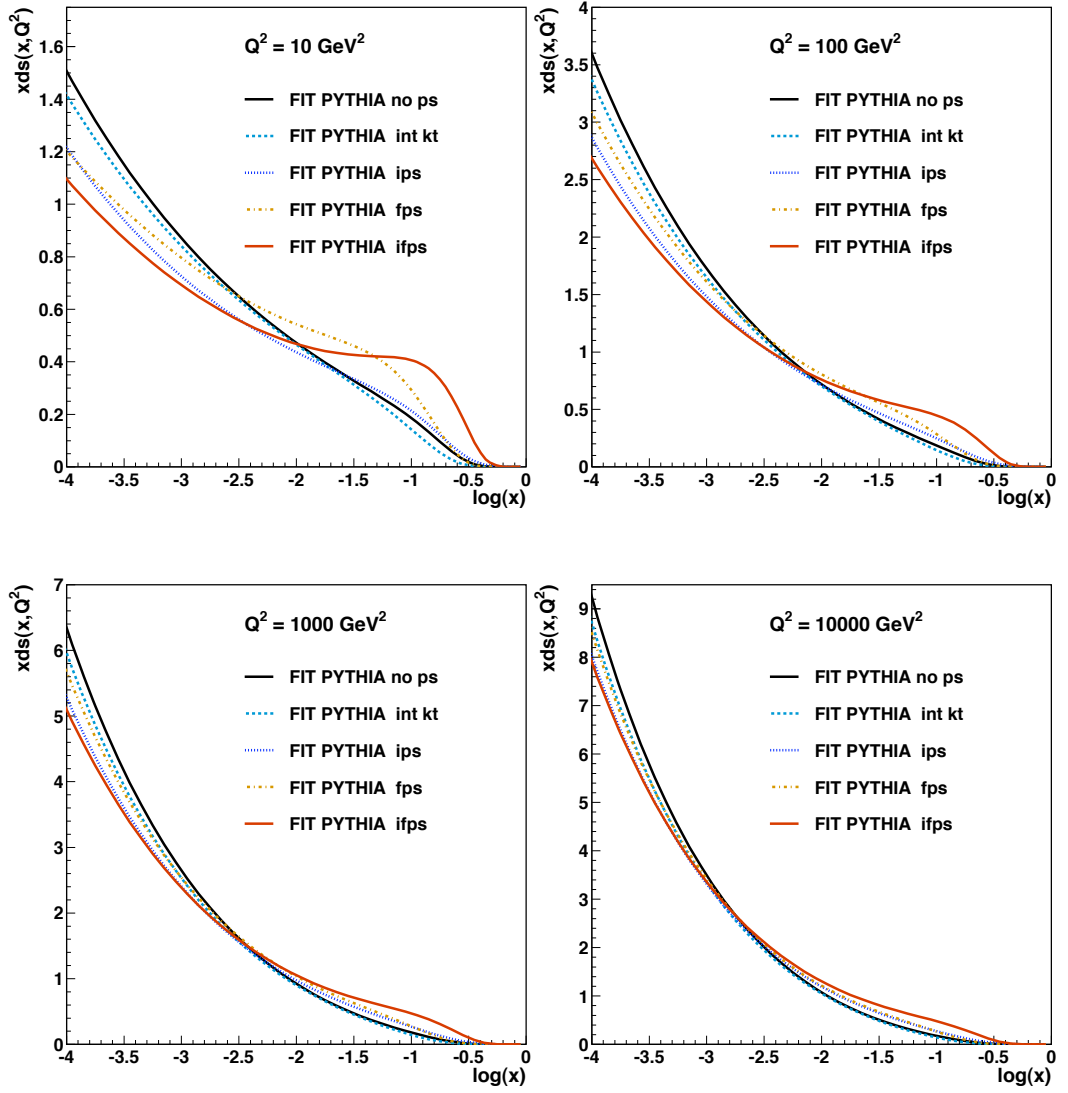


Figure 9.13: Down sea quark PDF as a function of x from k_{\perp} -ordered PYTHIA, fitted with no parton shower, no parton shower + intrinsic k_{\perp} , initial state parton shower, final state parton shower and the combination of initial and final state parton shower, at $Q^2 = 10 \text{ GeV}^2$, 100 GeV^2 , 1000 GeV^2 and 10000 GeV^2 .

9.3.2 Q^2 -ordered parton showers

Fig. 9.14 shows the gluon, fig. 9.15 the up sea quark and fig. 9.16 the down sea quark PDFs as a function of x resulting from the fit using PYTHIA for $Q^2 = 10, 100, 1000$ and 10000 GeV^2 . The parton shower were Q^2 -ordered.

In the x region below 10^{-2} the ik_{\perp} -PDFs, ips-PDFs, fps-PDFs and ifps-PDFs of gluons, up sea quarks and down sea quarks decrease in comparison to the respective nops-PDFs. The order (going from lowest decrease to highest decrease compared to the nops-PDFs) is hereby ips-PDFs, ik_{\perp} -PDFs, ifps-PDFs and fps-PDFs.

In the x region above 10^{-2} the fps-PDF increases in comparison to the nops-PDF of the gluons, the up sea quarks and the down sea quarks. The ifps-PDFs increase less than the fps-PDFs in comparison to the nops-PDFs. Even smaller is the increase of the ips-PDFs compared to the nops-PDFs.

The order (going from highest increase to lowest increase compared to the no parton shower PDF) is fps-PDFs, ifps-PDFs and ips-PDFs for gluons, up sea quarks and down sea quarks.

For increasing Q^2 , all PDFs show a clear reduction of the decrease observed for the inclusion of parton showers and intrinsic k_{\perp} .

The parameter values resulting from the fit are given in the tables B.3.

Interpretation

As in sec. 9.3.1 the observed changes in the PDFs can be explained by the change of x due to the virtualities gained in the parton shower.

The decrease at low x is stronger for PDFs obtained from fits with k_{\perp} -ordered parton shower than for PDFs obtained from fits with Q^2 -ordered parton shower. Looking at this in more detail shows, that the PDFs obtained from fits with k_{\perp} -ordered parton shower and from fits with Q^2 -ordered parton shower show a similar decrease of the fps-PDFs. However is the decrease for ips-PDFs in case of k_{\perp} -ordered parton showers stronger than that for the fps-PDFs, while it is weaker than that for the fps-PDFs in case of Q^2 -ordered parton shower.

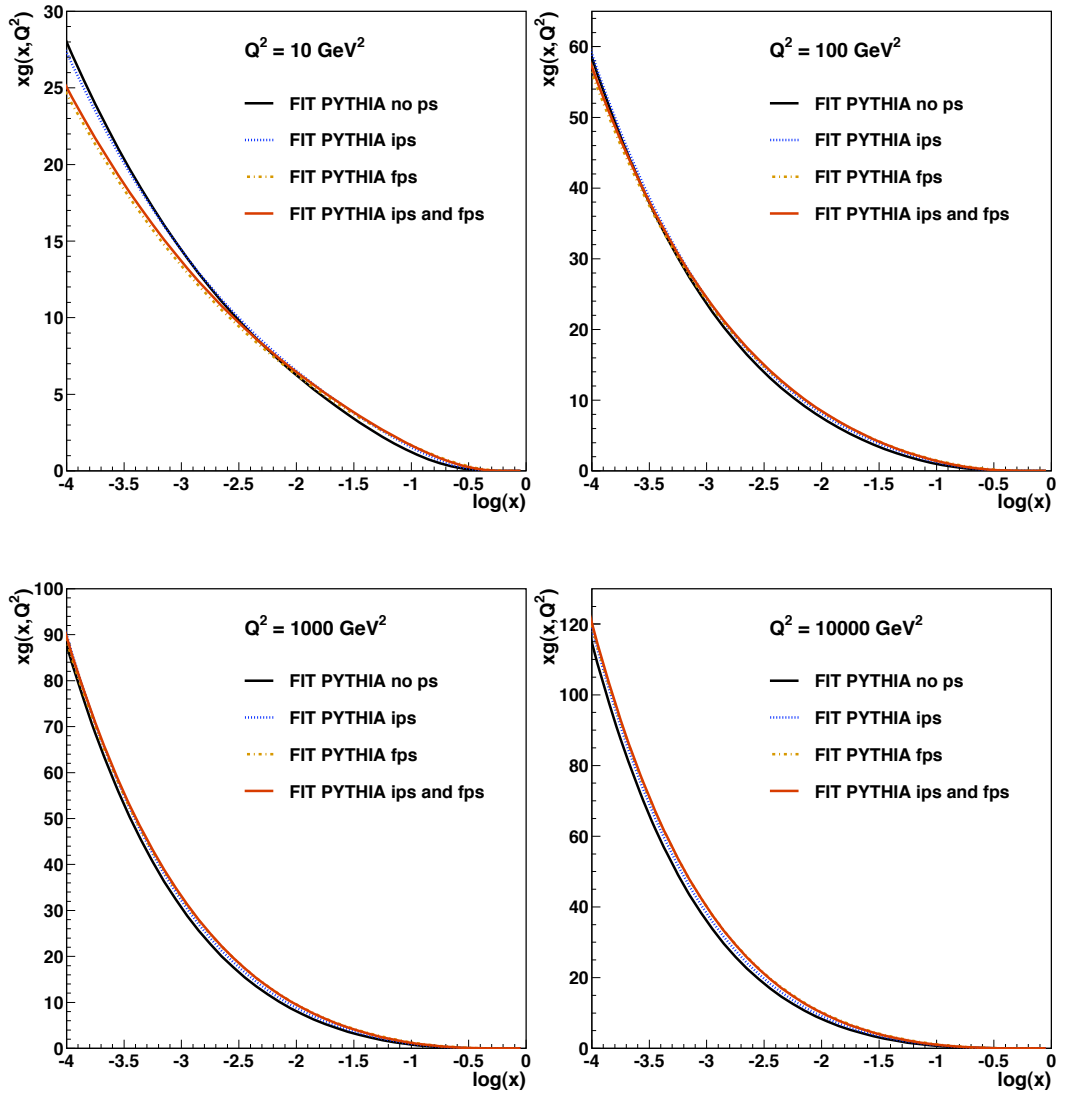


Figure 9.14: Gluon PDF as a function of x from Q^2 -ordered PYTHIA, fitted with no parton shower, initial state parton shower, final state parton shower and the combination of initial and final state parton shower, at $Q^2 = 10 \text{ GeV}^2$, 100 GeV^2 , 1000 GeV^2 and 10000 GeV^2 .

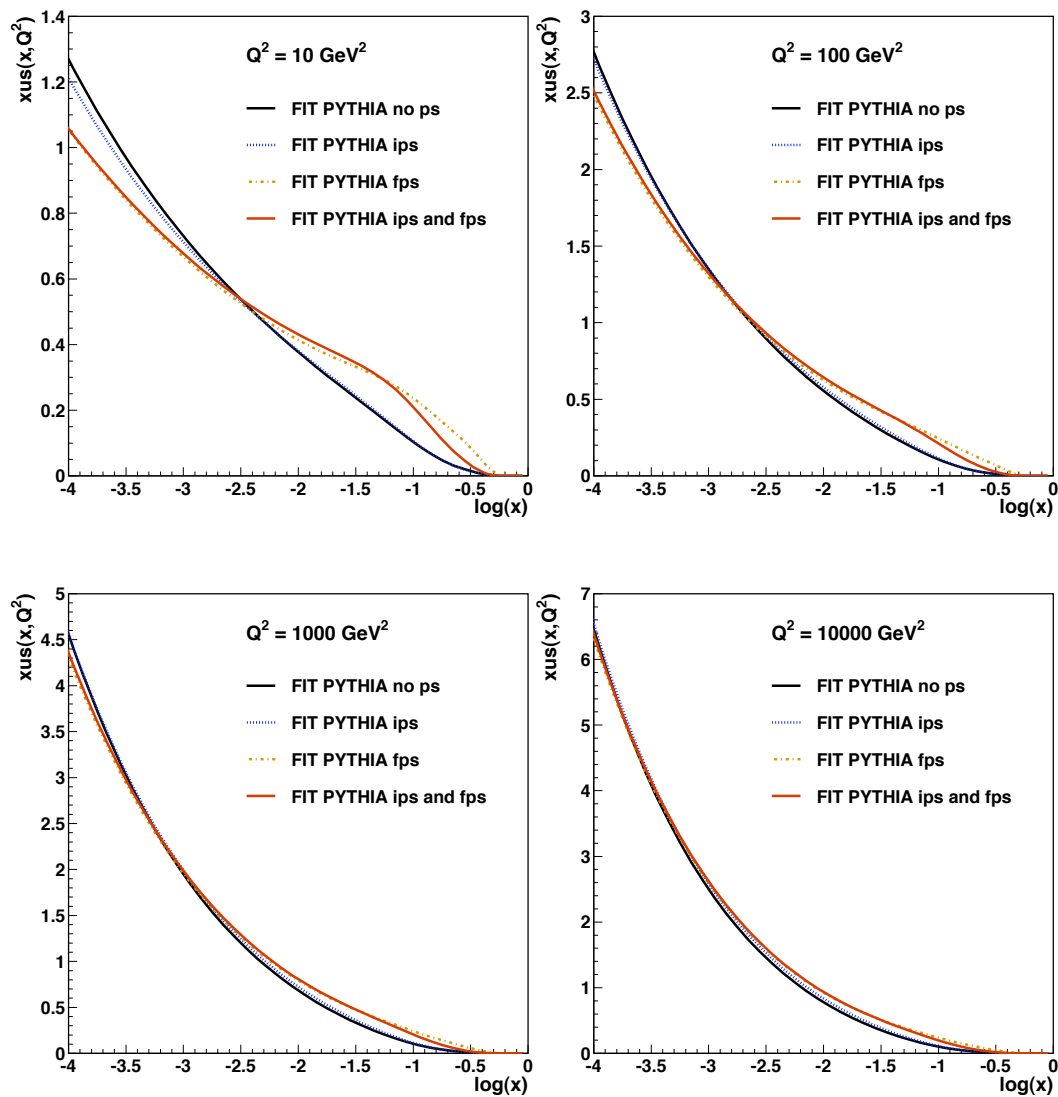


Figure 9.15: Up sea quark PDF as a function of x from Q^2 -ordered PYTHIA, fitted with no parton shower, initial state parton shower, final state parton shower and the combination of initial and final state parton shower, at $Q^2 = 10 \text{ GeV}^2$, 100 GeV^2 , 1000 GeV^2 and 10000 GeV^2 .

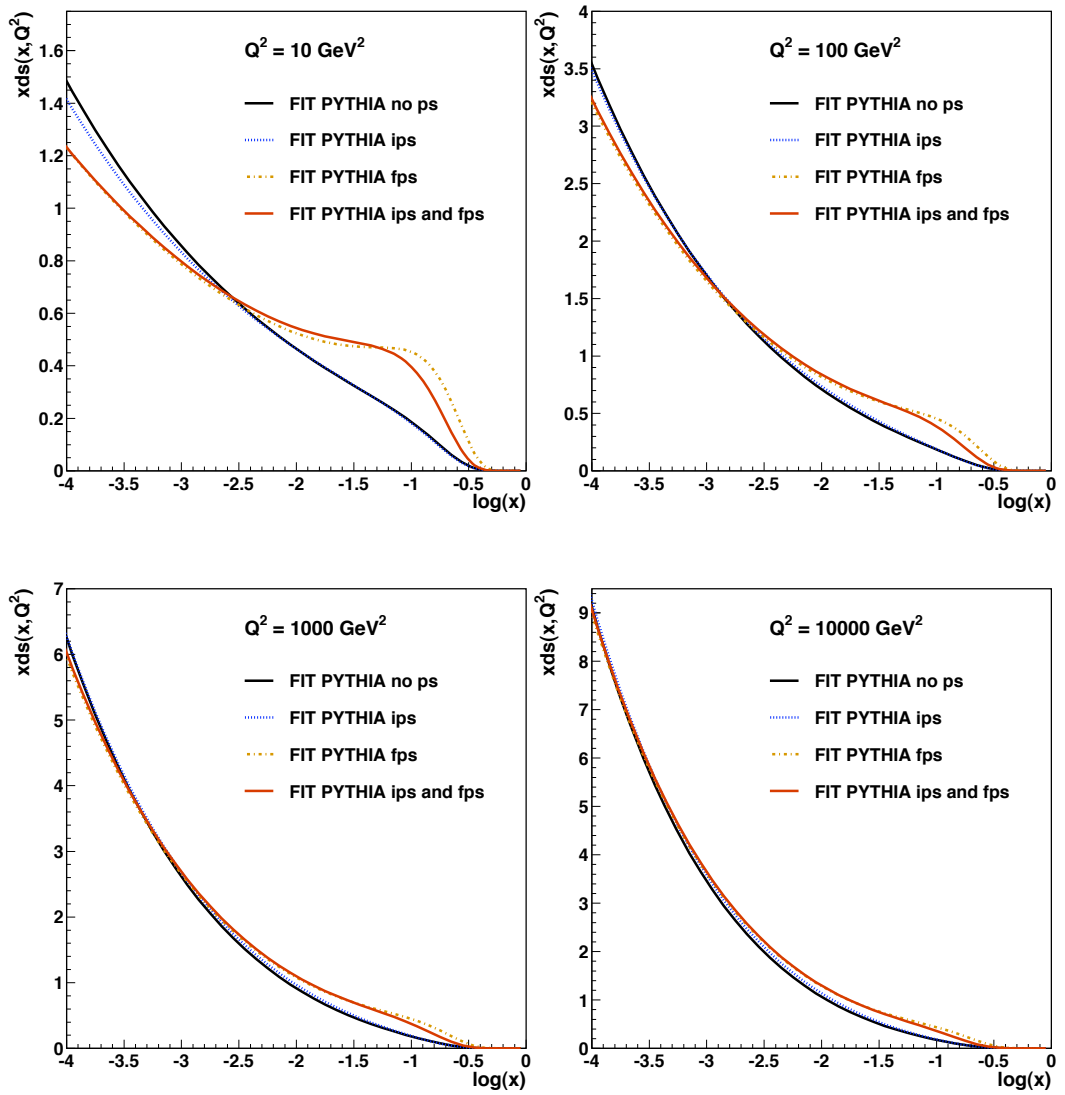


Figure 9.16: Down sea quark PDF as a function of x from Q^2 -ordered PYTHIA, fitted with no parton shower, initial state parton shower, final state parton shower and the combination of initial and final state parton shower, at $Q^2 = 10 \text{ GeV}^2$, 100 GeV^2 , 1000 GeV^2 and 10000 GeV^2 .

9.3.3 Comparison with NLO PDF set

Fig. 9.18 shows the gluon PDF as a function of x from the LO set CTEQ6L, the NLO set CTE6.1M (+ error band) and the ones resulting from the fit using PYTHIA with a Q^2 -ordered and with a k_{\perp} -ordered combination of initial and final state parton shower for $Q^2 = 10, 100, 1000$ and 10000 GeV^2 .

The highest PDF values are given by the CTEQ6L set. The next lower values are given in the lower x region by the PDF from the PYTHIA fit with Q^2 -ordered combination of initial and final state parton shower. Even lower are the PDF values given by the PDF obtained by the PYTHIA fit with k_{\perp} -ordered combination of initial and final state parton shower. The lowest PDF values are given by CTEQ6.1M.

At $Q^2 = 10 \text{ GeV}^2$ the k_{\perp} -ordered PDF lies from $x \approx 3 \cdot 10^{-3}$ on in the error band of CTEQ6.1M. For $Q^2 = 100, 1000$ and 10000 GeV^2 it lies completely outside the error band at all.

The Q^2 -ordered PDF lies outside the error band of CTEQ6.1M for all Q^2 .

The higher the Q^2 , the closer the PDFs resulting from a fit, using a Monte Carlo event generator, come to the CTEQ6L PDF set.

Interpretation

In an NLO matrix element the first radiative correction is - other than in an LO matrix element - included. By that, a radiation, that can reduce the value of x and has in an LO calculation to be included in the PDF, is now part of the matrix element (see fig. 9.17). This has the effect, that in an NLO PDF less radiations are taken into account in the PDFs and therefore the probability for low x is smaller, than in an LO calculation. This is why CTEQ6.1M is lower at small x than CTEQ6L.

The plots show that the PDFs resulting from a fit, using a Monte Carlo event generator, are closer to the NLO PDF CTEQ6.1 M, than to the CTEQ6L PDF. PDFs resulting from a fit including k_{\perp} -ordered parton shower are closer to the NLO PDFs, than PDFs resulting from a fit including Q^2 -ordered parton shower do.

The observation that the PDFs resulting from a fit using a Monte Carlo event generator come closer to the CTEQ6L PDF with increasing Q^2 can

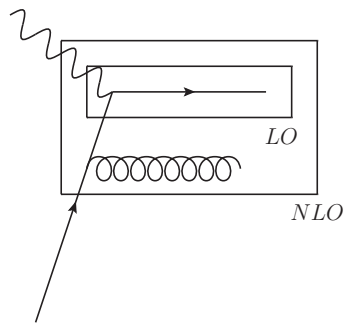


Figure 9.17: In contrast to an LO matrix element, the NLO matrix element contains the first radiation

again be explained by eq. (9.5). The higher Q^2 is compared to p^2 and m^2 the less significant becomes the difference between eq. (9.2) and eq. (9.5).

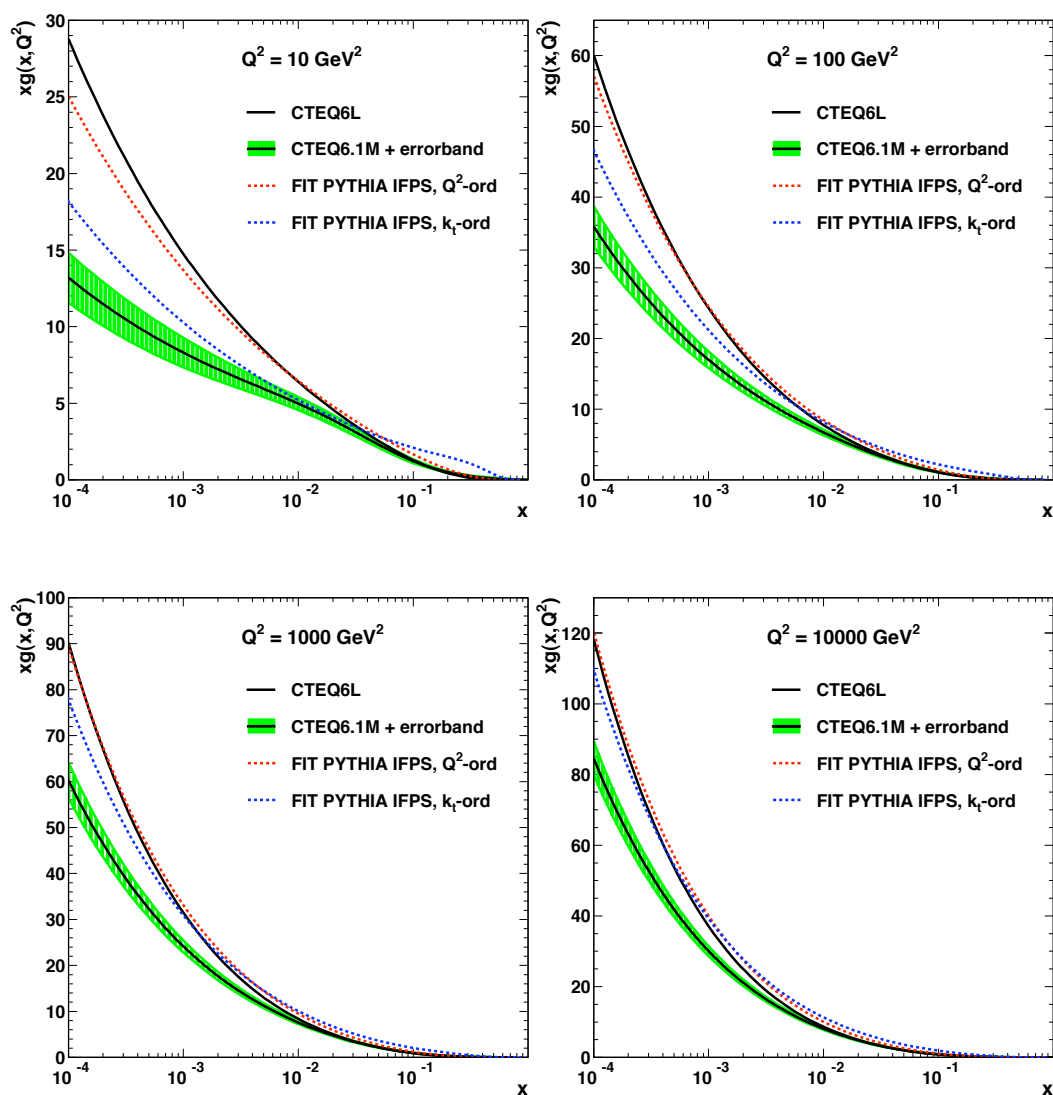


Figure 9.18: Gluon PDF as a function of x from Q^2 - and k_\perp -ordered PYTHIA, fitted with the combination of initial and final state parton shower, compared to LO PDF set CTEQ6L and NLO PDF set CTEQ5M at $Q^2 = 10 \text{ GeV}^2, 100 \text{ GeV}^2, 1000 \text{ GeV}^2$ and 10000 GeV^2 .

9.3.4 Comparison of χ^2

In table 9.1 the values of χ^2 for the comparison of F_2 data with the predictions from RAPGAP and PYTHIA running with different PDF sets are listed. For each kind of parton shower and intrinsic k_\perp respectively, χ^2 is given for using the PDF set CTEQ6L and the PDF set obtained by the fits. In all cases an improvement of χ^2 , and therefore the data description, resulted from a fit.

A slight improvement of χ^2 can be observed in RAPGAP when the PDFs resulting from a fit are used instead of CTEQ6L. The improvement is in all cases comparable and the χ^2 for the fitted PDFs is in all four cases a factor of approx. 1.2-1.8 lower than the χ^2 for CTEQ6L.

In the case of PYTHIA, there is a striking difference between the χ^2 values for using CTEQ6L and the PDF resulting from the fit, when parton shower are switched on.

In the k_\perp -ordered case, with the combination of initial and final state parton shower, the χ^2 value for using the PDF from the fit is approx. a factor of 12 lower than the χ^2 value for using CTEQ6L. This factor is approx. 7 when the combination of initial and final state parton shower is used in PYTHIA with Q^2 -ordered parton shower and only approx. 1.3 when this kind of parton shower is used in RAPGAP.

This indicates that the inclusion of parton shower in a Monte Carlo event generator like PYTHIA requires modified PDFs to describe data properly.

| generator and ordering | parton shower and PDF set | | | | | | | | | | | |
|---------------------------|---------------------------|--------|-------------------|--------|----------------|--------|-----------------|--------|-----------------------|--------|--------|--------|
| | no p.s. | | intr. k_{\perp} | | in. state p.s. | | fin. state p.s. | | in. + fin. state p.s. | | | |
| | CTEQ6L | fitted | CTEQ6L | fitted | CTEQ6L | fitted | CTEQ6L | fitted | CTEQ6L | fitted | CTEQ6L | fitted |
| PYTHIA, k_{\perp} -ord. | 139.12 | 105.03 | 283.34 | 157.96 | 1162.28 | 211.52 | 773.66 | 112.68 | 2180.42 | 182.40 | | |
| PYTHIA, Q^2 -ord. | 97.61 | 77.18 | - | - | 268.46 | 113.81 | 740.65 | 145.09 | 920.11 | 134.73 | | |
| RAPGAP | 106.29 | 89.81 | - | - | 134.97 | 74.06 | 108.00 | 77.78 | 113.72 | 87.52 | | |

Table 9.1: Values of χ^2 for RAPGAP and the two parton shower orderings of PYTHIA running with no parton shower, intrinsic k_{\perp} , initial state parton shower, final state parton shower and the combination of initial and final state parton shower. The χ^2 is calculated when the Monte Carlo event generators were running with the CTEQ6L PDF set (left subdivision) and with the fitted PDF (right subdivision).

Chapter 10

Conclusions and Outlook

The parton showers, added in Monte Carlo event generators, change normally the x distribution of the hard interacting parton and therefore the PDFs. Only the Monte Carlo event generators RAPGAP and LEPTO provide an exception.

The influence of parton shower on the PDFs was investigated in this work. Predictions from the Monte Carlo event generators RAPGAP and PYTHIA were fitted to F_2 data as measured at HERA. The gluon and sea quark PDFs were determined.

Three different scenarios were investigated: PDFs were fitted using RAPGAP and PYTHIA without parton shower, PDFs were fitted using RAPGAP with parton shower and PDFs were fitted using PYTHIA with a) k_{\perp} -ordered and b) Q^2 -ordered parton shower.

The PDFs resulting from the fits were compared with PDF sets, obtained by other numerical methods.

The PDFs obtained using RAPGAP and PYTHIA without parton shower agree reasonably well with the reference PDFs. This shows the consistency of the method, since Monte Carlo event generators without parton showers and the semi-analytical methods used to determine the reference set proceed similar.

Using RAPGAP with parton shower for the fit leads to PDFs that again agree with the reference PDFs. This is consistent, since RAPGAP does not change the x distribution of the hard interacting parton and Q^2 when parton shower are used and thus there is no influence on the PDFs. This thesis shows for the first time, that RAPGAP works properly in this aspect.

The PDFs obtained using PYTHIA including k_{\perp} -ordered parton shower show significant changes of the resulting PDFs compared to the reference PDFs.

The inclusion of parton shower in the generation process leads to a decrease of the resulting PDFs in the lower x region and an increase in the higher x region.

The PDFs obtained using PYTHIA including Q^2 -ordered parton shower show similar but smaller effects.

It is interesting to note that after including parton shower, the PDFs obtained by a fit using a Monte Carlo event generator become more similar to the NLO PDFs.

We conclude: that the method of using Monte Carlo event generators in fits to determine PDFs can be applied successfully, and that including parton shower lead in general to significant changes of the PDFs.

A modified PDF set for PYTHIA (PDF4MC) was determined in this thesis.

After this proof of concept, one can now extend this study to include the latest combined HERA data on inclusive structure functions, to obtain a precise set of PDF4MC.

Since the details of the parton shower treatment are important, this study has to be extended to determine PDF4MC also for HERWIG and other Monte Carlo event generators.

The PDF4MC will be a important input to the consistent physics simulation at all high energy colliders and especially at the LHC.

Bibliography

- [1] C. Adloff et al. Deep-inelastic inclusive e p scattering at low x and a determination of $\alpha(s)$. *Eur. Phys. J.*, C21:33–61, 2001.
- [2] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nuclear Physics B*, 126(2):298 – 318, 1977.
- [3] Mats Bengtsson and Torbjorn Sjostrand. A Comparative Study of Coherent and Noncoherent Parton Shower Evolution. *Nucl. Phys.*, B289:810, 1987.
- [4] Volker Blobel and Erich Lohrmann. *Statistische und numerische Methoden der Datenanalyse*. Teubner, 1998.
- [5] I. Borozan, N. Brook, J. Bromley, A. Buniatian, J. Butterworth, T. Carli, T. Ebert, J. Elmsheuser, G. Grindhammer, O. Gutsche, M. Hayes, E. Heaphy, D. Jansen, H. Jung, D. Kant, G. Knies, M. Kuhlen, L. Lönnblad, G. Nellen, P. van Mechelen, R. Mohr, A. Moraes, K. Peters, R. Poeschl, K. Sedlak, S. Soeldner-Rembold, K. Rabbertz, H. Rick, M. Riveline, F.-P. Schilling, Th. Schoerner, A. Solano, C. Targett-Adams, R. Taylor, A. Valkarova, B. Waugh, and T. Wengler. *HZTool A Package for Monte Carlo Generator – Data Comparisons at SPS, HERA, TeVatron, LEP and LHC (CVS version v3-004)*.
- [6] M. Botje, M. Klein, and C. Pascaud. Future precision measurements of $F_2(x, Q^2)$, $\alpha(s)(Q^2)$ and $xg(x, Q^2)$ at HERA. 1996.
- [7] R. Brock et al. Uncertainties of parton distribution functions and their implications on physical predictions. 2000.
- [8] Robin Devenish and Amanda Cooper-Sarkar. *Deep Inelastic Scattering*. Oxford University Press, 2004.

-
- [9] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and $e^+ e^-$ Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov. Phys. JETP*, 46:641–653, 1977.
- [10] V. N. Gribov and L. N. Lipatov. Deep inelastic $e p$ scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15:438–450, 1972.
- [11] V. N. Gribov and L. N. Lipatov. $e^+ e^-$ pair annihilation and deep inelastic $e p$ scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15:675–684, 1972.
- [12] J.D. Griffith. *An Introduction into Elementary Particle Physics*. Wiley&Sons, 2nd edition, 2008.
- [13] F. Halzen and A.D. Martin. *Quarks and Leptons, An Introductory Course in Modern Particle Physics*. John Wiley&Sons, 1984.
- [14] <http://projects.hepforge.org/lhapdf/>. *The Les Houches Accord PDF Interface*.
- [15] G. Ingelman, A. Edin, and J. Rathsman. *LEPTO 6.5 - A Monte Carlo Generator for Deep Inelastic Lepton-Nucleon Scattering*.
- [16] F. James. *Minuit - Function Minimization and Error Analysis*. CERN - Networks and Computing Division, 94.1 edition.
- [17] Hannes Jung. *The RAPGAP Monte Carlo for Deep Inelastic Scattering*. DESY, Notkestr.85, 22603 Hamburg, Germany, 3.1 edition.
- [18] L. N. Lipatov. The parton model and perturbation theory. *Sov. J. Nucl. Phys.*, 20:94–102, 1975.
- [19] J. Pumplin et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07:012, 2002.
- [20] Jun John Sakurai. *Modern Quantum Mechanics*. Addison Wesley, revised edition, 1994.
- [21] Peter Schmüser. *Feynman-Graphen und Eichtheorien für Experimentalphysiker*. Springer, 1995.
- [22] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. *PYTHIA 6.4, Physics and Manual*.

- [23] D. Stump et al. Uncertainties of predictions from parton distribution functions. 1. The Lagrange multiplier method. *Phys. Rev.*, D65:014012, 2002.
- [24] Deniz Sunar. *Measurement of $K^{*\pm}(892)$ Production in Deep Inelastic ep Scattering with the H1 Detector at HERA*. PhD thesis, UNIVERSITEIT ANTWERPEN, 2009.

Appendix A

From the solution of the dirac equation to the two-body cross section

A.1 Perturbative solution of the Dirac equation

The dirac equation in an electromagnetic field is:

$$(i\gamma^\mu\partial_\mu - m)\psi(x) = -e\gamma^\mu A_\mu(x)\psi(x) \quad (\text{A.1})$$

with the the γ -matrices

$$\begin{aligned} \gamma^0 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \\ \gamma^2 &= \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \end{aligned}$$

the covariant derivative

$$\partial^\mu = \left(\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z} \right) = \left(\frac{\partial}{\partial t}, -\vec{\nabla} \right), \quad (\text{A.2})$$

and the four vector potential

$$A_\mu = (\Phi, \vec{A}) \quad (\text{A.3})$$

A derivation of the Dirac equation and this specific form of it can be found in [21] or [13].

To focus now on the solution of (A.1) using a Greens function one uses the ansatz:

$$(i\gamma^\mu\partial_\mu - m)K(x, x') = \delta^4(x - x') \quad (\text{A.4})$$

where $K(x, x')$ is the desired Greens function. It is a 4x4 matrix and depends only on the difference $(x - x')$. Once $K(x, x')$ is found, a solution of (A.1) is:

$$\psi(x) = -e \int d^4x' K(x - x')\gamma^\mu A_\mu(x')\psi(x') \quad (\text{A.5})$$

since

$$\begin{aligned} (i\gamma^\mu\partial_\mu - m)\psi(x) &= -e \int d^4x' \underbrace{(i\gamma^\mu\partial_\mu - m)K(x - x')}_{\delta^4(x-x')} \gamma^\mu A_\mu(x')\psi(x') \\ &= -e\gamma^\mu A_\mu(x)\psi(x) \end{aligned} \quad (\text{A.6})$$

Any solution of the homogenous Dirac equation (that is a plane wave) can then be added

$$(i\gamma^\mu\partial_\mu - m)\Phi(x) = 0.$$

to obtain in the end an integral equation:

$$\psi(x) = \Phi(x) - e \int d^4x' K(x - x')\gamma^\mu A_\mu(x')\psi(x'). \quad (\text{A.7})$$

In contrast to (A.1) this equation has the advantage to be solvable iteratively by an expansion in the coupling constant α_e (this is justified by the smallness of $\alpha_e \approx \frac{1}{137}$). In this sense the second term of (A.7) is treated as a small perturbation. In the 0th-order approximation one can then write:

$$\psi^{(0)}(x) = \Phi(x). \quad (\text{A.8})$$

Inserting this into (A.7) then gives then the 1st-order term

$$\begin{aligned} \psi^{(1)}(x) &= \Phi(x) - e \int d^4x' K(x - x')\gamma^\mu A_\mu(x')\psi^{(0)}(x') \\ &= \Phi(x) - e \int d^4x' K(x - x')\gamma^\mu A_\mu(x')\Phi(x') \end{aligned} \quad (\text{A.9})$$

Inserting this term again into (A.7) then gives the 2nd-order term and so on.

A.2 Electron propagator

The calculation of $K(x - x')$ is shown in ([21]) with two results. Firstly that the Fourier transform of $K(x - x')$ is the *electron propagator* and secondly that the time evolution of a solution of the Dirac equation is given by:

$$\begin{aligned}\Phi(x) &= \Phi(t, \vec{x}) = i \int d^3x' K(x - x') \gamma^0 \Phi(t', \vec{x}') \\ \bar{\Phi}(x') &= \bar{\Phi}(t', \vec{x}') = i \int d^3x \bar{\Phi}(t, \vec{x}) \gamma^0 K(x - x') \\ &= i \int d^3x \Phi^\dagger(t, \vec{x}) K(x - x')\end{aligned}\quad (\text{A.10})$$

for $t' < t$ and with $\bar{\Phi} := \Phi^\dagger \gamma^0$.

A.3 Photon propagator

In the fermion-fermion or fermion-hadron scattering process the electromagnetic potential A^μ is present. A^μ can be obtained by solving the wave equation:

$$\square A^\mu(x) = \left(\frac{\partial^2}{\partial t^2} - \nabla^2 \right) A^\mu(x) = e j^\mu(x) \quad (\text{A.11})$$

where $e j^\mu(x) = e \bar{\psi}_f(x) \gamma^\mu \psi_i(x)$ is the four-vector current produced by the target and A^μ satisfies the Lorentz condition $\partial_\mu A^\mu = 0$.

One solves this equation again by using a Greens function satisfying:

$$\square D^{\mu\nu}(x - x') = g^{\mu\nu} \delta^4(x - x'). \quad (\text{A.12})$$

Then

$$A^\mu(x) = e \int d^4x' D^{\mu\nu}(x - x') j_\nu(x') \quad (\text{A.13})$$

To find $D^{\mu\nu}(x - x')$ one calculates its fourier transform and applies the Laplace operator to it.

$$\begin{aligned}\square D^{\mu\nu}(x - x') &= \square \int \frac{d^4q}{(2\pi)^4} \tilde{D}^{\mu\nu}(q) e^{-iq(x-x')} \\ &= \int \frac{d^4q}{(2\pi)^4} \tilde{D}^{\mu\nu}(q) (-q^2) e^{-iq(x-x')} \stackrel{!}{=} g^{\mu\nu} \delta^4(x - x')\end{aligned}\quad (\text{A.14})$$

Recalling that $\delta^4(x - x') = \int \frac{d^4q}{(2\pi)^4} e^{-iq(x-x')}$ yields:

$$\tilde{D}^{\mu\nu}(q) = \frac{-g^{\mu\nu}}{q^2 + i\epsilon} \quad (\text{A.15})$$

The ϵ is added to avoid divergencies and can be set to zero later. $\tilde{D}^{\mu\nu}(q)$ is the *photon propagator*.

A.4 Transition matrix

In an experimental setup as shown in fig. A.1 the incoming particle is described by a wave packet of momentum \vec{p}_i .

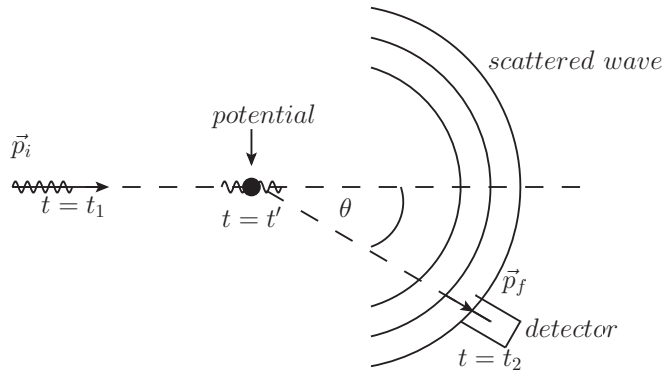


Figure A.1: Scattering of a wave packet from a potential

The wave packet can be replaced by a plane wave Φ_i , which is allowed, assuming that the electromagnetic potential is in a limited time-space-volume. At time $t = t_1$ the particle has yet to experience a potential. At time $t = t'$ it reaches the target and a spherical outgoing scattered wave ψ_{scat} is produced, which reaches the detector at $t = t_2$.

The detector is located at an angle θ to the trajectory of the incoming particle and covers a piece of solid angle $\Delta\Omega$. Therefore only the part of the scattered wave in the direction of the detector (parallel to momentum vector \vec{p}_f) is measured. Mathematically this means, that the scattered wave has to be expanded in plane waves and projected onto the wave Φ_f with momentum \vec{p}_f .

To calculate a matrix element for the transition $\Phi_i \rightarrow \Phi_f$ the transition matrix S is defined as:

$$\psi_{scat} = S \cdot \Phi_i. \quad (\text{A.16})$$

The transition $\Phi_i \rightarrow \Phi_f$ is then given by

$$S_{if} = \langle \Phi_f | S | \Phi_i \rangle = \int d^3x_2 \Phi_f^\dagger S \Phi_i = \int d^3x_2 \Phi_f^\dagger \psi_{scat} \quad (\text{A.17})$$

The 1st-order scattered wave can be rewritten according to (A.9) and inserted in (A.17), to give the 1st-order transition matrix:

$$S_{if}^{(1)} = -e \int d^4x' \int \underbrace{d^3x_2 \Phi_f^\dagger(x_2) K(x_2 - x')}_{(A.10) \Rightarrow \bar{\Phi}_f(x')} \gamma^\mu A_\mu(x') \Phi_i(x'). \quad (A.18)$$

This leads to the important result:

$$S_{if}^{(1)} = ie \int d^4x \bar{\Phi}_f(x') \gamma^\mu A_\mu(x') \Phi_i(x') \quad (A.19)$$

A.5 Fermion-fermion scattering

One is now in the position to calculate the matrix element for a fermion-fermion transition (fig. A.2) which will be used extensively later. With

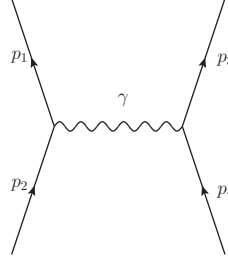


Figure A.2: Fermion-fermion scattering

$$\psi_i(x) = u(p_2)e^{-ip_2x} \quad \text{and} \quad \psi_f(x) = u(p_4)e^{-ip_4x}$$

one gets

$$ej^\mu(x) = e\bar{u}(p_4)\gamma^\mu u(p_2)e^{i(p_4-p_2)x} \quad (A.20)$$

This and the photon propagator (A.15) are now inserted into (A.13) to yield:

$$A^\mu(x) = e \int d^4x' \int \frac{d^4q}{(2\pi)^4} \frac{-g^{\mu\nu}}{q^2 + i\epsilon} e^{i(p_4-p_2+q)x'} e^{-iqx} \bar{u}(p_4)\gamma_\nu u(p_2) \quad (A.21)$$

where

$$q^2 = (p_3 - p_1)^2$$

is the photon's four momentum.

Integrating over x' gives a deltafunction $\delta^4(p_4 - p_2 + q)$. To then obtain the matrix element we have to calculate (A.19) using

$$\Phi_i(x) = u(p_1)e^{-ip_1x} \quad \text{and} \quad \bar{\Phi}_f(x) = \bar{u}(p_3)e^{-ip_3x}.$$

The x -integration in (A.19) gives again a deltafunction $(2\pi)^4 \delta^4(p_3 - p_1 - q)$ resulting in:

$$S_{if}^{(1)} = ie^2 \int d^4q \delta^4(p_4 - p_2 + q) \delta^4(p_3 - p_1 - q) (2\pi)^4 \cdot \bar{u}(p_3) \gamma^\mu u(p_1) \frac{-g^{\mu\nu}}{q^2 + i\epsilon} \bar{u}(p_4) \gamma_\nu u(p_2) \quad (\text{A.22})$$

This is the matrix element for fermion-fermion scattering (from now on I will leave the superscript "(1)"). One can directly write it down, using the Feynman rules (see sec. 2.1).

Integrating over the photons four momentum q then yields:

$$S_{fi} = \mathcal{M} \cdot (2\pi)^4 \delta^4(p_3 + p_4 - p_1 - p_2) \quad (\text{A.23})$$

with

$$\mathcal{M} := -ie^2 \bar{u}(p_3) \gamma_\mu u(p_1) \cdot \frac{1}{q^2} \cdot \bar{u}(p_4) \gamma^\mu u(p_2). \quad (\text{A.24})$$

A.6 Two-body cross section

The differential cross section of a two-body scattering process is calculated in the *center of mass system* (CMS). Then:

$$p_1 = (E_1, \vec{p}), \quad p_2 = (E_2, -\vec{p}), \quad p_3 = (E_3, \vec{p}'), \quad p_4 = (E_4, -\vec{p}')$$

and

$$s = (p_1 + p_2)^2 \quad (\text{A.25})$$

Starting with eq. (2.6) an integration over all quantities that are not observed has to be carried out. For the p_4 -dependent part in d_{Lips} the delta function can be used

$$\int d^3p_4 \delta^3(\vec{p}_3 + \vec{p}_4 - \underbrace{\vec{p}_1 - \vec{p}_2}_{=0 \text{ in CMS}}) \frac{\delta(E_3 + E_4 - E_1 - E_2)}{E_4} = \frac{\delta(E_3 + E_4 - E_1 - E_2)}{E_4}$$

where now

$$E_3 = \sqrt{|\vec{p}'|^2 + m_3^2}, \quad E_4 = \sqrt{|\vec{p}'|^2 + m_4^2}$$

This can be used to rewrite $\frac{d^3p_3}{E_3}$ in d_{Lips} . Since now $\frac{dE_3}{dE_4} = \frac{1}{2} \frac{1}{E_3} \frac{d|\vec{p}'|^2}{dE_4}$ and $|\vec{p}'|^2 = E_4^2 - m_4^2$ it is clear that $E_3 dE_3 = E_4 dE_4$. By defining $s' = (E_3 + E_4)^2$

one can write $dE_3 = \frac{E_4}{\sqrt{s'}} d\sqrt{s'}$. Using further spherical coordinates one can write:

$$d^3p_3 = |\vec{p}'|^2 d|\vec{p}'| d\Omega = |\vec{p}'| E_3 dE_3 d\Omega \Rightarrow \frac{d^3p_3}{E_3} = |\vec{p}'| dE_3 d\Omega = |\vec{p}'| E_4 \frac{d\sqrt{s'}}{\sqrt{s'}} d\Omega$$

which then gives:

$$dL_{ips} = (2\pi)^4 \delta(\sqrt{s} - \sqrt{s'}) \frac{1}{4(2\pi)^6} \frac{|\vec{p}'|}{\sqrt{s'}} d\sqrt{s'} d\Omega \quad (\text{A.26})$$

Using eq. (2.4) the *flux factor* in the CMS takes the form

$$\text{flux factor} = 4(|\vec{p}|E_1 + |\vec{p}|E_2) = 4|\vec{p}|\sqrt{s} \quad (\text{A.27})$$

The integrated cross section can then be written

$$\begin{aligned} d\sigma &= |\mathcal{M}|^2 \int \frac{dL_{ips}}{\text{flux factor}} \\ &= \int \frac{1}{4\sqrt{s'}|\vec{p}|} |\mathcal{M}|^2 (2\pi)^4 \delta(\sqrt{s} - \sqrt{s'}) \frac{1}{4(2\pi)^6} \frac{|\vec{p}'|}{\sqrt{s'}} d\sqrt{s'} d\Omega \end{aligned}$$

Integrating over $\sqrt{s'}$ then produces the differential cross section in the CMS:

$$\frac{d\sigma}{d\Omega} = \frac{1}{(8\pi)^2 s} \frac{|\vec{p}'|}{|\vec{p}|} |\mathcal{M}|^2 \quad (\text{A.28})$$

By assuming massless particles (which gives $p_1^2 = p_2^2 = 0$ and therefore $|\vec{p}|^2 = E_1^2 = E_2^2$) one gets

$$s = (p_1 + p_2)^2 = 4|\vec{p}|^2. \quad (\text{A.29})$$

Using this and

$$\begin{aligned} Q^2 &= -q^2 \\ \Rightarrow dQ^2 &= -dq^2 = 2|\vec{p}'||\vec{p}'| d(-\cos\theta) = \frac{|\vec{p}'||\vec{p}'|}{\pi} d\Omega \end{aligned}$$

one can rewrite A.28 as

$$\boxed{\frac{d\sigma}{dQ^2} = \frac{1}{16\pi s^2} |\mathcal{M}|^2} \quad (\text{A.30})$$

Appendix B

Fitted parameters

| parameter | fitresult |
|-----------------------|-----------|
| A_g | 3.2839 |
| B_g | -0.067018 |
| C_g | 7.3155 |
| D_g | 4.4233 |
| E_g | 1.5341 |
| F_g | -1.5565 |
| $A_{\bar{u}+\bar{d}}$ | 0.11115 |
| $B_{\bar{u}+\bar{d}}$ | -0.25538 |
| $C_{\bar{u}+\bar{d}}$ | 8.8384 |
| $D_{\bar{u}+\bar{d}}$ | 2.6179 |
| $E_{\bar{u}+\bar{d}}$ | 3.2633 |
| $F_{\bar{u}+\bar{d}}$ | 0.61680 |

(a)

| parameter | fitresult |
|-----------------------|-----------|
| A_g | 3.4889 |
| B_g | -0.046194 |
| C_g | 7.4079 |
| D_g | 4.5011 |
| E_g | 1.5927 |
| F_g | -1.1885 |
| $A_{\bar{u}+\bar{d}}$ | 0.13077 |
| $B_{\bar{u}+\bar{d}}$ | -0.23656 |
| $C_{\bar{u}+\bar{d}}$ | 9.4767 |
| $D_{\bar{u}+\bar{d}}$ | 2.6531 |
| $E_{\bar{u}+\bar{d}}$ | 3.2818 |
| $F_{\bar{u}+\bar{d}}$ | 0.34017 |

(b)

| parameter | fitresult |
|-----------------------|-----------|
| A_g | 3.3832 |
| B_g | -0.057657 |
| C_g | 7.3339 |
| D_g | 4.4615 |
| E_g | 1.5707 |
| F_g | -1.3853 |
| $A_{\bar{u}+\bar{d}}$ | 0.13093 |
| $B_{\bar{u}+\bar{d}}$ | -0.23633 |
| $C_{\bar{u}+\bar{d}}$ | 8.8443 |
| $D_{\bar{u}+\bar{d}}$ | 2.8067 |
| $E_{\bar{u}+\bar{d}}$ | 3.2807 |
| $F_{\bar{u}+\bar{d}}$ | 0.54130 |

(c)

| parameter | fitresult |
|-----------------------|-----------|
| A_g | 3.3171 |
| B_g | -0.060359 |
| C_g | 7.3311 |
| D_g | 4.4401 |
| E_g | 1.5456 |
| F_g | -1.5331 |
| $A_{\bar{u}+\bar{d}}$ | 0.11856 |
| $B_{\bar{u}+\bar{d}}$ | -0.25159 |
| $C_{\bar{u}+\bar{d}}$ | 8.8692 |
| $D_{\bar{u}+\bar{d}}$ | 2.6820 |
| $E_{\bar{u}+\bar{d}}$ | 3.2544 |
| $F_{\bar{u}+\bar{d}}$ | 0.61406 |

(d)

Table B.1: Parameters resulting from RAPGAP-fit to F_2 data with no parton shower (B.1a), initial state parton shower (B.1b), final state parton shower (B.1c), combination of initial and final state parton shower (B.1d)

| parameter | fitresult |
|-----------------------|-----------|
| A_g | 3.3598 |
| B_g | -0.064154 |
| C_g | 7.3119 |
| D_g | 4.4325 |
| E_g | 1.5201 |
| F_g | -1.5090 |
| $A_{\bar{u}+\bar{d}}$ | 0.12249 |
| $B_{\bar{u}+\bar{d}}$ | -0.24418 |
| $C_{\bar{u}+\bar{d}}$ | 8.9656 |
| $D_{\bar{u}+\bar{d}}$ | 2.6463 |
| $E_{\bar{u}+\bar{d}}$ | 3.2626 |
| $F_{\bar{u}+\bar{d}}$ | 0.61009 |

(a)

| parameter | fitresult |
|-----------------------|-----------|
| A_g | 3.1576 |
| B_g | -0.054406 |
| C_g | 7.3274 |
| D_g | 4.4414 |
| E_g | 1.5526 |
| F_g | -1.1929 |
| $A_{\bar{u}+\bar{d}}$ | 0.14136 |
| $B_{\bar{u}+\bar{d}}$ | -0.22473 |
| $C_{\bar{u}+\bar{d}}$ | 12.665 |
| $D_{\bar{u}+\bar{d}}$ | 3.1397 |
| $E_{\bar{u}+\bar{d}}$ | 3.3422 |
| $F_{\bar{u}+\bar{d}}$ | 0.57039 |

(b)

| parameter | fitresult |
|-----------------------|-----------|
| A_g | 2.7635 |
| B_g | -0.032760 |
| C_g | 4.4053 |
| D_g | 7.2568 |
| E_g | 1.7199 |
| F_g | -1.4004 |
| $A_{\bar{u}+\bar{d}}$ | 0.11551 |
| $B_{\bar{u}+\bar{d}}$ | -0.25040 |
| $C_{\bar{u}+\bar{d}}$ | 8.9459 |
| $D_{\bar{u}+\bar{d}}$ | 2.6658 |
| $E_{\bar{u}+\bar{d}}$ | 3.2631 |
| $F_{\bar{u}+\bar{d}}$ | 0.61747 |

(c)

| parameter | fitresult |
|-----------------------|-----------|
| A_g | 3.1556 |
| B_g | -0.017320 |
| C_g | 7.9125 |
| D_g | 8.2652 |
| E_g | 2.0625 |
| F_g | -1.1173 |
| $A_{\bar{u}+\bar{d}}$ | 0.20710 |
| $B_{\bar{u}+\bar{d}}$ | -0.16073 |
| $C_{\bar{u}+\bar{d}}$ | 13.361 |
| $D_{\bar{u}+\bar{d}}$ | 3.4031 |
| $E_{\bar{u}+\bar{d}}$ | 3.5453 |
| $F_{\bar{u}+\bar{d}}$ | 0.89884 |

(d)

| parameter | fitresult |
|-----------------------|------------|
| A_g | 1.6289 |
| B_g | -0.0036833 |
| C_g | 7.9146 |
| D_g | 13.882 |
| E_g | 2.0677 |
| F_g | -1.1393 |
| $A_{\bar{u}+\bar{d}}$ | 0.15044 |
| $B_{\bar{u}+\bar{d}}$ | -0.20945 |
| $C_{\bar{u}+\bar{d}}$ | 9.2984 |
| $D_{\bar{u}+\bar{d}}$ | 7.8029 |
| $E_{\bar{u}+\bar{d}}$ | 3.3656 |
| $F_{\bar{u}+\bar{d}}$ | 0.59902 |

(e)

Table B.2: Parameters resulting from PYTHIA-fit to F_2 data with k_\perp -ordered no parton shower (B.2a), intrinsic k_\perp (B.2b), initial state parton shower (B.2c), final state parton shower (B.2d), combination of initial and final state parton shower (B.2e)

| parameter | fitresult |
|-----------------------|-----------|
| A_g | 3.3547 |
| B_g | -0.060047 |
| C_g | 7.3303 |
| D_g | 4.4328 |
| E_g | 1.5458 |
| F_g | -1.4425 |
| $A_{\bar{u}+\bar{d}}$ | 0.11745 |
| $B_{\bar{u}+\bar{d}}$ | -0.24792 |
| $C_{\bar{u}+\bar{d}}$ | 8.8514 |
| $D_{\bar{u}+\bar{d}}$ | 2.6425 |
| $E_{\bar{u}+\bar{d}}$ | 3.2720 |
| $F_{\bar{u}+\bar{d}}$ | 0.62133 |

(a)

| parameter | fitresult |
|-----------------------|-----------|
| A_g | 3.3992 |
| B_g | -0.043116 |
| C_g | 7.7987 |
| D_g | 4.7065 |
| E_g | 0.91859 |
| F_g | -1.3448 |
| $A_{\bar{u}+\bar{d}}$ | 0.11615 |
| $B_{\bar{u}+\bar{d}}$ | -0.24348 |
| $C_{\bar{u}+\bar{d}}$ | 8.8760 |
| $D_{\bar{u}+\bar{d}}$ | 2.6438 |
| $E_{\bar{u}+\bar{d}}$ | 3.2759 |
| $F_{\bar{u}+\bar{d}}$ | 0.57173 |

(b)

| parameter | fitresult |
|-----------------------|-----------|
| A_g | 3.1527 |
| B_g | -0.019139 |
| C_g | 7.8689 |
| D_g | 8.4510 |
| E_g | 2.3052 |
| F_g | -0.94060 |
| $A_{\bar{u}+\bar{d}}$ | 0.16136 |
| $B_{\bar{u}+\bar{d}}$ | -0.19058 |
| $C_{\bar{u}+\bar{d}}$ | 11.716 |
| $D_{\bar{u}+\bar{d}}$ | 9.0130 |
| $E_{\bar{u}+\bar{d}}$ | 3.4104 |
| $F_{\bar{u}+\bar{d}}$ | 0.79989 |

(c)

| parameter | fitresult |
|-----------------------|------------|
| A_g | 3.5461 |
| B_g | -0.0041712 |
| C_g | 7.5625 |
| D_g | 6.9383 |
| E_g | 1.8212 |
| F_g | -1.1046 |
| $A_{\bar{u}+\bar{d}}$ | 0.16237 |
| $B_{\bar{u}+\bar{d}}$ | -0.18806 |
| $C_{\bar{u}+\bar{d}}$ | 12.529 |
| $D_{\bar{u}+\bar{d}}$ | 5.5388 |
| $E_{\bar{u}+\bar{d}}$ | 3.5246 |
| $F_{\bar{u}+\bar{d}}$ | 0.99711 |

(d)

Table B.3: Parameters resulting from PYTHIA-fit to F_2 data with Q^2 -ordered no parton shower (B.3a), initial state parton shower (B.3b), final state parton shower (B.3c), combination of initial and final state parton shower (B.3d)

Danksagung

Ich möchte meinem Betreuer vor Ort, Dr. Hannes Jung, danken, für seine Geduld und seine Unterstützung bei dieser Arbeit. Prof. Fumagalli und Prof. Schrader möchte ich danken, für ihre Befürwortung dieser Arbeit, ohne die, diese nie zustande gekommen wäre. I want to thank Albert Knutsson for bringing me back on track and the entire group for a great time. Ich danke Matthi, für seinen Rat und den Spaß, den wir haben, und Vinzenz, für unsere unerschütterliche Freundschaft.

Vor allem aber möchte ich meinen Eltern, meiner Schwester, meinem Schwager und meinen Nichten danken, für die Freude, die sie mir bringen, und dafür, dass sie für mich da sind!

Erklärung:

Ich versichere, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Ort, Datum

Federico von Samson – Himmelstjerna