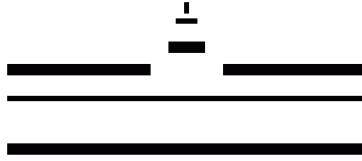


University of Münster  
– Westfälische Wilhelms-Universität Münster –



Institute of Theoretical Physics



Master's thesis

---

# Parton Distribution Functions from Deep Inelastic Scattering

---

<b>handed in by:</b>	PIT DUWENTÄSTER
<b>handed in on:</b>	AUGUST 20, 2019
<b>Primary Supervisor:</b>	PD DR. KAROL KOVAŘÍK
<b>Secondary Supervisor:</b>	PROF. DR. MICHAEL KLASSEN

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Scattering processes</b>	<b>4</b>
2.1	Electron-muon scattering . . . . .	4
2.2	Elastic electron-proton scattering . . . . .	6
2.3	Deep inelastic scattering at tree level . . . . .	8
2.4	$Z$ -boson exchange and interference . . . . .	11
2.5	QED vertex correction . . . . .	13
2.6	The parton model . . . . .	16
2.6.1	Basic statements and consequences of the parton model . . . . .	16
2.6.2	The quark structure of nucleons . . . . .	20
2.7	The QCD improved parton model . . . . .	24
2.7.1	Real emission of gluons . . . . .	26
2.7.2	DGLAP equations . . . . .	31
2.7.3	Scale dependence . . . . .	33
<b>3</b>	<b>Modelling of data</b>	<b>35</b>
3.1	Least squares fitting . . . . .	35
3.2	Nonlinear least squares fitting . . . . .	37
3.2.1	Levenberg-Marquardt method . . . . .	39
3.2.2	Implementation in C++ . . . . .	40
3.2.3	$K$ factors . . . . .	43
3.2.4	Examples . . . . .	44
3.3	The Hessian formalism for uncertainties . . . . .	46
3.3.1	Parameterization of PDFs and the effective $\chi^2$ . . . . .	46
3.3.2	The Hessian formalism . . . . .	47
3.3.3	Estimation of $T$ and $t$ . . . . .	51
3.3.4	Implementation in C++ . . . . .	51
3.3.5	Fit improvement . . . . .	54
3.3.6	Examples . . . . .	55

<b>4</b>	<b>Data analysis</b>	<b>58</b>
4.1	Preparation of data . . . . .	58
4.2	PDF fitting . . . . .	60
4.2.1	nCTEQ parameterization . . . . .	60
4.2.2	PDF fits . . . . .	62
4.2.3	Comparison with data . . . . .	67
<b>5</b>	<b>Summary and outlook</b>	<b>70</b>
<b>A</b>	<b>Appendix</b>	<b>71</b>
A.1	Summaries and derivations . . . . .	71
A.1.1	Regularization . . . . .	71
A.1.2	$D$ -dimensional spherical coordinates . . . . .	72
A.1.3	Wick rotation and generic integrals . . . . .	73
A.1.4	Mellin transform and Mellin convolution . . . . .	75
A.1.5	Plus distribution . . . . .	75
A.1.6	Feynman rules . . . . .	76
A.1.7	Group theory . . . . .	79
A.2	Frequently used relations and identities . . . . .	81
A.2.1	$\gamma$ -matrices . . . . .	81
A.2.2	Traces . . . . .	82
A.2.3	Feynman parameterization . . . . .	82
A.2.4	Gamma- and Beta-functions . . . . .	83
A.2.5	Lab frame kinematics and kinematic variables . . . . .	83
A.2.6	Photon polarization sum . . . . .	84
	<b>Bibliography</b>	<b>86</b>

# 1 Introduction

Scattering experiments have been crucial to the development of particle physics as we know it today. From the Geiger-Marsden / Rutherford experiments [1] in 1911 leading to vast improvements in our understanding of atoms to modern collider experiments at facilities like the LHC where the existence of the Higgs-boson as the last missing particle of the Standard Model was confirmed in 2012 [2], [3], scattering experiments continue to provide invaluable insight into the nature of our universe.

In the late 1960s Richard Feynman proposed that hadrons are built up of point-like constituents called partons [4], which are now known to be gluons and the quarks of Gell-Mann's quark model [5]. The parton distribution functions (PDFs) describing the structure of hadrons cannot be calculated in perturbation theory and therefore need to be calculated from lattice gauge theory or extracted from experimental data. This thesis focuses on the latter approach. The PDFs are a key component for predictions of any high energy scattering process involving hadrons and with the frequent releases of new experimental data comes a constant motivation to improve the PDFs [6]. The aim of this thesis is therefore to provide a general overview of the ingredients that go into the determination of PDFs as well as in-depth discussion of an explicit example for each major aspect. This includes precision calculation of the theory, fitting algorithms, and the estimation of uncertainties. The thesis concludes with the determination of PDFs from actual experimental data.

The thesis begins with the examination of various scattering processes leading up to deep inelastic scattering (DIS) [7]–[10], which was instrumental in the historical development of the parton model and is still part of any modern global analysis. From there on the leading-order parton model is explored before calculating the QCD corrections at next-to-leading order (NLO) to DIS in the parton model. The result of the NLO-QCD correction are the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equations [11]–[14], which allow the perturbative calculations of the PDFs' dependence on the momentum scale  $Q^2$ .

The second part of the thesis is dedicated to the numerical techniques necessary for the determination of PDFs and their uncertainties. The Levenberg-Marquardt algorithm [15], [16] is chosen as a method of fitting arbitrary analytic functions to data and the required changes for fitting PDFs, which are generally not of analytically known form, are discussed. To allow the quantitative determination of the fit's uncertainties, the Hessian formalism [17], [18] is introduced and then tested on some illustrative fits. Both methods are implemented in C++ and integrated into the nCTEQ codebase, which provides the precision computations of the theory necessary for PDF fitting.

Finally the results of the previous sections are combined with data from the NMC [19] and BCDMS [20], [21] experiments to perform multiple PDF fits, which are then compared among each other and have some validation checks performed on them.



## 2 Scattering processes

As mentioned in the introduction, scattering processes are an essential part in both the history of particle physics and also in modern particle physics. Results from inelastic scattering experiments were the reason for the proposition of the parton model, which is the basis of this thesis. For this reason, we will first derive the equations describing the differential cross section of various scattering-processes and -models, including elastic electron-muon scattering, elastic electron-proton, and lastly inelastic electron-proton scattering. Subsequently, their relations among each other and to the parton model will be explained.

### 2.1 Electron-muon scattering

We begin the theoretical investigation of scattering processes with electron-muon scattering (or more generally scattering of two distinguishable point-like fermions) as the most simple model. The invariant amplitude  $\mathcal{M}$  is given by the diagram<sup>1</sup>

$$i\mathcal{M} = \begin{array}{c} e^- \\ \swarrow k \\ \bullet \\ \downarrow q \\ \bullet \\ \swarrow p \\ \mu^- \end{array} \begin{array}{c} e^- \\ \searrow k' \\ \bullet \\ \downarrow q \\ \bullet \\ \searrow p' \\ \mu^- \end{array} = -e^2 \bar{u}(k') \gamma^\mu u(k) \frac{1}{q^2} e^2 \bar{u}(p') \gamma^\mu u(p). \quad (1)$$

To calculate the unpolarized cross section we need the spin-averaged square of  $\mathcal{M}$ , which can be expressed as

$$|\overline{\mathcal{M}}|^2 = \frac{e^4}{q^4} L_e^{\mu\nu} L_{\mu\nu}^{\text{muon}}, \quad (2)$$

with the electron's leptonic tensor

$$L_e^{\mu\nu} = \frac{1}{2} \sum_{e \text{ spins}} [\bar{u}(k') \gamma^\mu u(k)] [\bar{u}(k') \gamma^\nu u(k)]. \quad (3)$$

This tensor can be expressed in terms of explicit matrix elements as

$$L_e^{\mu\nu} = \frac{1}{2} \sum_{s'} \bar{u}_\alpha^{s'}(k') \gamma_{\alpha\beta}^\mu \sum_s u_\beta^s(k) \bar{u}_\rho^s(k') \gamma_{\rho\sigma}^\nu u_\sigma^{s'}(k). \quad (4)$$

Now we can freely move  $u_\sigma^{s'}(k)$  and use  $\sum_s u^s(k) \bar{u}^s(k) = \not{k} + m$  to get

$$L_e^{\mu\nu} = \frac{1}{2} \text{Tr}((\not{k}' + m) \gamma^\mu (\not{k} + m) \gamma^\nu), \quad (5)$$

---

<sup>1</sup>All Feynman diagrams are created using the L<sup>A</sup>T<sub>E</sub>X package TikZ-Feynman [22]

which we can rewrite using the trace relations from section A.2.2. This gives

$$L_e^{\mu\nu} = 2(k'^\mu k^\nu + k^\mu k'^\nu - k' \cdot k g^{\mu\nu} + m^2 g^{\mu\nu}). \quad (6)$$

The leptonic tensor of the muon is mostly the same, aside from replacing the electron mass  $m$  by the muon mass  $M$ . Carrying out the Lorentz contraction of the two tensors results in

$$|\overline{\mathcal{M}}|^2 = 8 \frac{e^4}{q^4} [(k' \cdot p')(k \cdot p) + (k' \cdot p)(k \cdot p') - (k' \cdot k)M^2 - (p' \cdot p)m^2 + 2m^2 M^2]. \quad (7)$$

If we assume the muon at rest and neglect the electron mass we can simplify this to

$$|\overline{\mathcal{M}}|^2 = 8 \frac{e^4}{q^4} \left[ -\frac{1}{2} q^2 (k \cdot p - k' \cdot p) + 2(k' \cdot p)(k \cdot p) + \frac{1}{2} q^2 M^2 \right]. \quad (8)$$

Using the lab frame kinematics from section A.2.5 we get

$$|\overline{\mathcal{M}}|^2 = 8 \frac{e^4}{q^4} 2M^2 E E' \left[ \cos^2 \frac{\theta}{2} - \frac{q^2}{2M^2} \sin^2 \frac{\theta}{2} \right]. \quad (9)$$

This can be used to calculate the differential cross section  $d\sigma$  by considering the initial flux and the phase space of the outgoing particles. The differential cross section then reads

$$d\sigma = \frac{1}{2E2M} \frac{|\overline{\mathcal{M}}|^2}{4\pi^2} \delta^{(4)}(p + k - p' - k') \frac{d^3 k'}{2E'} \frac{d^3 p'}{2p'_0}. \quad (10)$$

Using the relation  $\frac{d^3 k'}{2E'} = \frac{1}{2} E' dE' d\Omega$  and plugging (9) into (10) we get

$$d\sigma = \frac{1}{2\pi^2} M E'^2 \frac{e^4}{q^4} \left( \cos^2 \frac{\theta}{2} + \frac{q^2}{2M^2} \sin^2 \frac{\theta}{2} \right) dE' d\Omega \frac{d^3 p'}{2p'_0} \delta^{(4)}(p + q - p'). \quad (11)$$

Further, we want to get rid of the differential  $p'$  by integrating

$$\int \frac{d^3 p'}{2p'_0} \delta^{(4)}(p + k - p' - k') = \frac{1}{2M} \delta^{(4)} \left( \nu + \frac{q^2}{2M} \right), \quad (12)$$

so that we get

$$\frac{d\sigma}{dE' d\Omega} = \frac{(2\alpha E')^2}{q^4} \left( \cos^2 \frac{\theta}{2} - \frac{q^2}{2M^2} \sin^2 \frac{\theta}{2} \right) \delta \left( \nu + \frac{q^2}{2M} \right). \quad (13)$$

Finally, by using the lab frame kinematics  $q^2 = -4EE' \sin^2 \frac{\theta}{2}$  and  $\nu = E - E'$  and performing the integration over  $E'$ , we arrive at

$$\left( \frac{d\sigma}{d\Omega} \right)_{\text{lab}} = \frac{\alpha^2}{4E^2 \sin^4 \frac{\theta}{2}} \frac{E'}{E} \left( \cos^2 \frac{\theta}{2} - \frac{q^2}{2M^2} \sin^2 \frac{\theta}{2} \right). \quad (14)$$

This formula describes the scattering of two point-like fermions. Since the proton is also a

fermion, it would be reasonable to expect scattering on a proton to be described by this equation. However, deviations seen in experimental measurements from this cross section implied that the proton has some form of substructure. Later on the return of the electron-proton cross section to this form in the limit of very high energies will be what leads us to the parton model.

## 2.2 Elastic electron-proton scattering

The cross section of scattering processes involving an electron and an arbitrary charge distribution instead of a point-like particle can be described by the cross section of a point-like particle times a form factor  $|F(q)|^2$ :

$$\frac{d\sigma}{d\Omega} = \left( \frac{d\sigma}{d\Omega} \right)_{\text{point}} |F(q)|^2. \quad (15)$$

It can be shown that for a static target the form factor is equivalent to the Fourier transform of the charge distribution  $\rho$

$$F(\mathbf{q}) = \int d^3x \rho(\mathbf{x}) e^{i\mathbf{q} \cdot \mathbf{x}}. \quad (16)$$

Expanding this for small momentum transfer  $|\mathbf{q}|$  and assuming spherical symmetry, this results in

$$\begin{aligned} F(\mathbf{q}) &= \int d^3x \rho(\mathbf{x}) \left( 1 + i\mathbf{q} \cdot \mathbf{x} - \frac{(\mathbf{q} \cdot \mathbf{x})^2}{2} + \mathcal{O}(\mathbf{q}^3) \right) \\ &= 1 - \frac{1}{6} |\mathbf{q}|^2 \langle r^2 \rangle + \mathcal{O}(\mathbf{q}^4). \end{aligned} \quad (17)$$

This gives us the mean square radius of the charge distribution and therefore a measure of the proton's size. The soft photon with small  $|\mathbf{q}|$  cannot resolve any further structure aside from the size of the charge distribution.

The case of the proton is a bit more complicated than a charge distribution, however, since the proton also has a magnetic moment and it will recoil during the scattering process. To describe the process properly one needs to redo the same calculation as for the electron-muon scattering while replacing the  $\gamma^\mu$  matrices in the muon's leptonic tensor (3) by

$$\gamma^\mu \xrightarrow{\text{proton}} \Gamma^\mu \equiv F_1(q^2) \gamma^\mu + \frac{\kappa}{2M} F_2(q^2) i\sigma^{\mu\nu} q_\nu, \quad (18)$$

which is the most general Lorentz four vector that can be constructed from  $q, p, p'$  and the  $\gamma$ -matrices ( $\kappa$  is the anomalous magnetic moment). Terms involving  $\gamma^5$  violate parity and are therefore irrelevant to electromagnetic interactions. Repeating the calculation from chapter 2.1 with this change then gives us the Rosenbluth formula

$$\left( \frac{d\sigma}{d\Omega} \right)_{\text{lab}} = \frac{\alpha^2}{4E^2 \sin^4 \frac{\theta}{2}} \frac{E'}{E} \left( \left( F_1^2 - \frac{\kappa^2 q^2}{4M^2} F_2^2 \right) \cos^2 \frac{\theta}{2} - \frac{q^2}{2M^2} (F_1 + \kappa F_2)^2 \sin^2 \frac{\theta}{2} \right). \quad (19)$$

The Rosenbluth formula looks very similar to (14), which can be restored as a special case when  $F_1 = 1$  and either  $\kappa = 0$  or  $F_2 = 0$ . The two form factors can be determined by measuring the cross section as a function of  $\theta$  and  $q^2$ . In practice it is often easier to write them in terms of

$$\begin{aligned} G_E &\equiv F_1 + \frac{\kappa q^2}{4M^2} F_2 \\ G_M &\equiv F_1 + \kappa F_2, \end{aligned} \quad (20)$$

so that there are no interference terms  $G_E G_M$  in the formula:

$$\left( \frac{d\sigma}{d\Omega} \right)_{\text{lab}} = \frac{\alpha^2}{4E^2 \sin^4 \frac{\theta}{2}} \frac{E'}{E} \left( \frac{G_E^2 + \tau G_M^2}{1 + \tau} \cos^2 \frac{\theta}{2} - 2\tau G_M^2 \sin^2 \frac{\theta}{2} \right), \quad \text{with } \tau \equiv -\frac{q^2}{4M^2}. \quad (21)$$

The form factors  $G_E(q^2)$  and  $G_M(q^2)$  are related to the charge and magnetic moment distributions of the proton, but due to the proton's recoil they cannot be interpreted as the Fourier transform of the distributions. However, approximating the Fourier transform of the charge distribution by the form factors is an approximation that holds in the case of  $|\mathbf{q}|^2 \ll M^2$ . Looking at the data shown in 2.1 shows that both  $G_E$  and  $G_M$  have a  $q^2$  dependence of the form

$$G_E = \frac{G_M}{\mu} = \left( 1 - \frac{q^2}{0.71 \text{ GeV}^2} \right)^{-2}. \quad (22)$$

Using equation (17) this gives us a mean square proton charge radius of

$$\langle r^2 \rangle = 6 \left( \frac{dG_E(q^2)}{dq^2} \right)_{q^2=0} = (0.81 \cdot 10^{-15} \text{ m})^2. \quad (23)$$

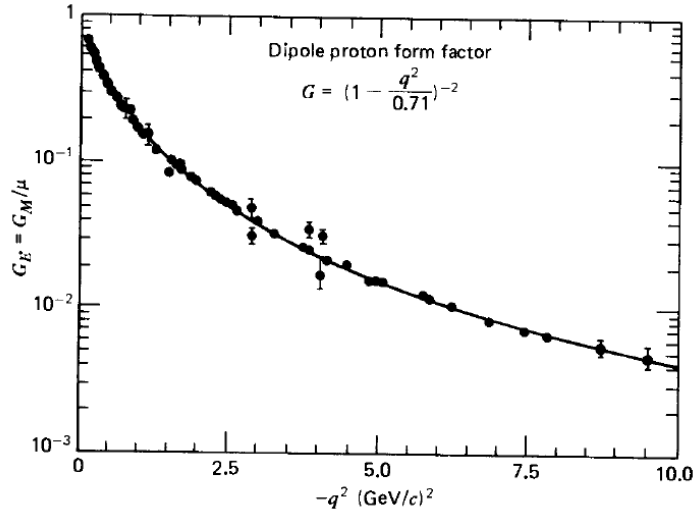


Figure 2.1:  $G_E$  as a function of  $Q^2$ . Data are from the Stanford Linear Accelerator, figure from [23].

### 2.3 Deep inelastic scattering at tree level

The result of the previous section was a measure for the size of the proton. In this section, we want to gain further insight about the details of its structure. This inevitably leads to Deep Inelastic Scattering (DIS) experiments, which have historically been one of the most important steps in gaining a better understanding of the structure of hadrons [7]–[9]. Even today they are used in precision measurements of the structure of nucleons and more recently even full atomic nuclei [6], [10]. The key difference between elastic and inelastic scattering is, that the outgoing particles of the inelastic scattering are not the same as the incoming ones and generally it is not known which particles are part of the result. The scattered electron transfers some energy to the proton, which can then perform a variety of reactions including breaking up into its constituents. In the case of the proton breaking up one speaks of deep inelastic scattering. Figure 2.2 schematically shows what happens in such a process. Another example of inelastic scattering is the excitation of  $\Delta$ -resonances, where the invariant mass  $W$  of the resulting particle is just slightly higher than the protons and it may decay into a proton and a  $\pi^0$  after some time. These resonances can be seen in figure 2.3 as the peaks in the cross section on the right of the elastic peak. The continuous tail of the cross section beyond those resonances shows the deep inelastic scattering region that we are interested in.

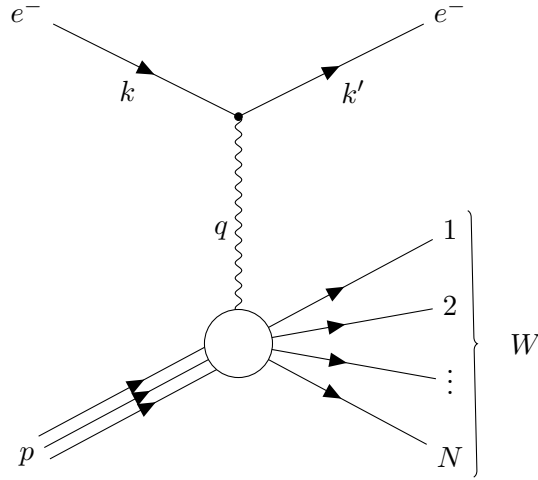


Figure 2.2: Lowest order diagram for deep inelastic scattering  $e^-p \rightarrow e^-X$ . The missing mass  $W$  is the sum of the energies of the final state particles in the bottom half of the diagram.

When calculating the cross section of this ( $e^-p \rightarrow e^-X$ )-process, we can make use of large portions of the previous ( $e^-\mu^- \rightarrow e^-\mu^-$ )-scattering process. The only thing that needs to be changed is that the leptonic tensor of the muon will be replaced by the hadronic tensor. Since we do not know anything about the final state particles we have to assume the most general form depending only on  $g^{\mu\nu}$ ,  $q$  and  $p$ :

$$W^{\mu\nu} = -W_1 g^{\mu\nu} + \frac{W_2}{M^2} p^\mu p^\nu + \left( \frac{i}{2M^2} \epsilon^{\mu\nu\kappa\lambda} p_\kappa q_\lambda W_3 \right) + \frac{q^\mu q^\nu}{M^2} W_4 + \frac{p^\mu q^\nu + q^\mu p^\nu}{M^2} W_5 \quad (24)$$

For now we can ignore the  $W_3$  term as it violates parity and is only relevant in electroweak calculations (we will come back to it in section 2.4). Taking into account the charge conservation

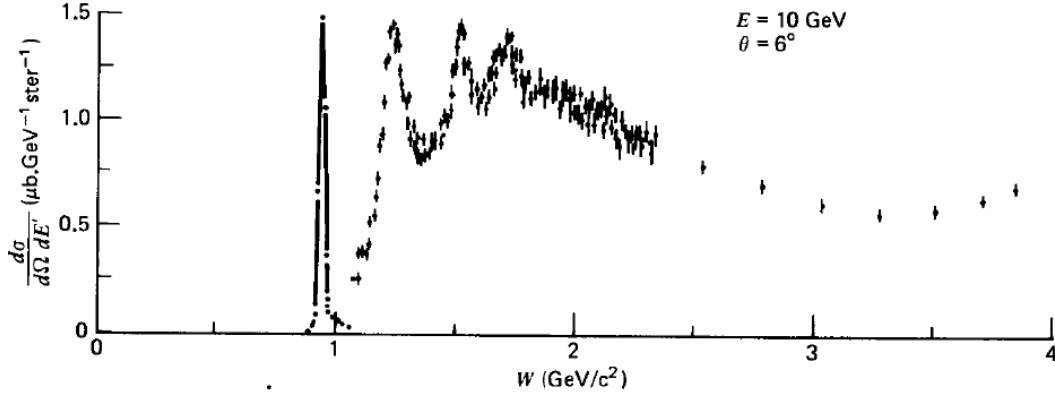


Figure 2.3: Cross section of electron-proton scattering as a function of missing mass  $W$ . The elastic peak has been scaled down by a factor of 8.5. Data are from the Stanford Linear Accelerator, the figure from [23].

$q_\mu W^{\mu\nu} = 0$  we can see that  $W_4$  and  $W_5$  can be expressed in terms of  $W_1$  and  $W_2$

$$W_4 = \frac{M^2}{q^2} W_1 + \left( \frac{p \cdot q}{q^2} \right)^2 W_2 \quad \text{and} \quad W_5 = -\frac{p \cdot q}{q^2} W_2. \quad (25)$$

Therefore the hadronic tensor can be expressed as

$$W^{\mu\nu} = W_1 \left( -g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2} \right) + \frac{W_2}{M^2} \left( p^\mu - \left( \frac{p \cdot q}{q^2} \right) q^\mu \right) \left( p^\nu - \left( \frac{p \cdot q}{q^2} \right) q^\nu \right). \quad (26)$$

Here,  $W_1$  and  $W_2$  are scalars that can only depend on the scalar variables at the vertex. We choose  $q^2$  and  $\nu = \frac{p \cdot q}{M}$  as our two independent variables. The square of the invariant mass <sup>2</sup>  $W^2$  can be written in terms of these variables as

$$W^2 = (p + q)^2 = p^2 + 2p \cdot q + q^2 = M^2 + 2M\nu + q^2. \quad (27)$$

Other common choices for the two independent variables are  $x = \frac{-q^2}{2p \cdot q} = \frac{-q^2}{2M\nu}$  and  $y = \frac{p \cdot q}{p \cdot k}$ . Given the tensors  $L_e^{\mu\nu}$  and  $W^{\mu\nu}$  it is possible to calculate their contraction. For the part involving  $W_1$  we get

$$\begin{aligned} L_e^{\mu\nu} W_{\mu\nu}^{(1)} &= 2 \left( k^\mu k'^\nu + k'^\mu k^\nu - (k \cdot k' - m^2) g^{\mu\nu} \right) W_1 \left( -g_{\mu\nu} + \frac{q_\mu q_\nu}{q^2} \right) \\ &= 2W_1 \left( k^\mu k'^\nu + k'^\mu k^\nu - (k \cdot k' - m^2) g^{\mu\nu} \right) (-g_{\mu\nu}) + (\text{terms} \propto q_\mu L^{\mu\nu} = 0) \\ &= 4W_1 k \cdot k'. \end{aligned} \quad (28)$$

---

<sup>2</sup> $W^2$  and the hadronic tensor  $W^{\mu\nu}$  have no direct relation, even though their similar names might imply otherwise.

The remaining part gives us

$$\begin{aligned}
L_e^{\mu\nu} W_{\mu\nu}^{(2)} &= 2 \left( k^\mu k'^\nu + k'^\mu k^\nu - (k \cdot k' - m^2) g^{\mu\nu} \right) \frac{W_2}{M^2} \left( p_\mu - \left( \frac{p \cdot q}{q^2} \right) q_\mu \right) \left( p_\nu - \left( \frac{p \cdot q}{q^2} \right) q_\nu \right) \\
&= 2 \frac{W_2}{M^2} \left( k^\mu k'^\nu + k'^\mu k^\nu - (k \cdot k' - m^2) g^{\mu\nu} \right) (p_\mu p_\nu) + (\text{terms} \propto q_\mu L^{\mu\nu} = 0) \\
&= 2W_2 \left( 2(p \cdot k)(p \cdot k') - M^2 k \cdot k' \right).
\end{aligned} \tag{29}$$

Combining the two parts (28) and (29) yields

$$L_e^{\mu\nu} W_{\mu\nu} = 4W_1 k \cdot k' + 2W_2 \left( 2(p \cdot k)(p \cdot k') - M^2 k \cdot k' \right). \tag{30}$$

In the lab frame:  $\left( k \cdot k' = 2EE' \sin^2 \frac{\theta}{2}, p \cdot k = EM, p \cdot k' = E'M \right)$

$$L_e^{\mu\nu} W_{\mu\nu} = 4EE' \left\{ W_2(\nu, q^2) \cos^2 \frac{\theta}{2} + 2W_1(\nu, q^2) \sin^2 \frac{\theta}{2} \right\} \tag{31}$$

This gives us the square of the invariant amplitude

$$|\overline{\mathcal{M}}|^2 = \frac{e^4}{q^4} 4\pi M (L^e)^{\mu\nu} W_{\mu\nu}. \tag{32}$$

The factor  $4\pi M$  comes from the convention for the normalization of  $W_{\mu\nu}$ . This finally allows us to calculate the cross section via

$$d\sigma = \frac{1}{4EM} |\overline{\mathcal{M}}|^2 \frac{d^3 k'}{2E'(2\pi)^3}, \quad \text{with} \quad \frac{d^3 k'}{2E'} = \frac{1}{2} E' dE' d\Omega. \tag{33}$$

This results in

$$\frac{d\sigma}{dE' d\Omega} = \frac{\alpha^2}{4E^2 \sin^4 \frac{\theta}{2}} \left\{ W_2(\nu, q^2) \cos^2 \frac{\theta}{2} + 2W_1(\nu, q^2) \sin^2 \frac{\theta}{2} \right\}. \tag{34}$$

We can compare this with other types of scattering that we have derived before:  $\frac{d\sigma}{dE' d\Omega} = \frac{4\alpha^2 E'^2}{q^4} \{ \dots \}$

- $e^- \mu^- \rightarrow e^- \mu^- (\text{elastic}) : \{ \dots \} = \left\{ \cos^2 \frac{\theta}{2} - \frac{q^2}{2m^2} \sin^2 \frac{\theta}{2} \right\} \delta \left( \nu + \frac{q^2}{2m} \right),$  (35)

- $e^- p^+ \rightarrow e^- p^+ (\text{elastic}) : \{ \dots \} = \left\{ \frac{G_E^2 + \tau G_M^2}{1 + \tau} \cos^2 \frac{\theta}{2} - 2\tau G_M^2 \sin^2 \frac{\theta}{2} \right\} \delta \left( \nu + \frac{q^2}{2m} \right),$  (36)

- $e^- p^+ \rightarrow e^- X : \{ \dots \} = \left\{ W_2(\nu, q^2) \cos^2 \frac{\theta}{2} - 2W_1(\nu, q^2) \sin^2 \frac{\theta}{2} \right\}.$  (37)

We have to include the  $\delta$ -functions in the first two to make sure that the scattering is actually elastic. As a check for self-consistency we can see that the cross section calculated for  $ep \rightarrow$

$eX$  can reproduce the results for both previous cases. However, since we have no way to determine the coefficients  $W_1(\nu, q^2)$  and  $W_2(\nu, q^2)$  from the theory, we need to measure them experimentally. The result of these measurements will then be leading to the parton model.

## 2.4 Z-boson exchange and interference

At very high energies, beyond  $Q^2 \equiv |q^2| \geq M_Z^2$ , we need to account for the coupling to the Z-boson in both the hadronic and the leptonic tensor. The coupling of the Z-boson to the fermions includes a parity violating part proportional to the  $\gamma^5$ -matrix. When the coupling is written as a matrix element it is different from the photons in the following way:

$$-iee_f \bar{u}(k') \gamma^\mu u(k) \longrightarrow -i \frac{g}{\cos(\theta_W)} \bar{u}(k') \gamma^\mu (C_L P_L + C_R P_R) u(k), \quad (38)$$

where  $g$  is the coupling constant of the  $SU(2)_L$  group in the electroweak Standard Model and  $\theta_W$  is the Weinberg mixing-angle. The coupling coefficients  $C_L$  and  $C_R$  can be written in terms of electroweak parameters  $I_f^3$  (isospin of the fermion) and  $e_f$  the fermion's charge (relative to a positron's charge). Alternatively one can parameterize it through axial and vector couplings  $a_v$  and  $g_v$ :

$$i \frac{g}{\cos(\theta_W)} \bar{u}(k') \gamma^\mu (C_L P_L + C_R P_R) u(k) = -i \frac{g}{\cos(\theta_W)} \bar{u}(k') \gamma^\mu (g_v - a_v \gamma^5) u(k), \quad (39)$$

so that

$$\begin{aligned} g_v &= C_L + C_R \quad \text{and} \quad a_v = C_L - C_R \\ \text{or equivalently} \quad C_L &= \frac{1}{2}(g_v + a_v) \quad \text{and} \quad C_R = \frac{1}{2}(g_v - a_v). \end{aligned} \quad (40)$$

Our goal is to get the cross section of the process in the following form:

$$\frac{d\sigma}{dE' d\Omega} = \frac{\alpha^2}{q^4} \frac{E'}{E} \sum_i \eta_i \tilde{L}_i^{\mu\nu} W_{\mu\nu}^i, \quad \text{with} \quad i \in \{\gamma\gamma, \gamma Z, Z\gamma, ZZ\}, \quad (41)$$

where we have a factor  $\eta$  that exists due to the different coupling constants between the lepton and the different exchange bosons:

$$\eta_{\gamma\gamma} = 1, \quad \eta_{\gamma Z} = \eta_{Z\gamma} = \frac{1}{4s_W^2 c_W^2} \frac{q^2}{q^2 - M_Z^2}, \quad \eta_{ZZ} = \frac{1}{16s_W^4 c_W^4} \frac{q^4}{(q^2 - M_Z^2)^2}. \quad (42)$$

We end up needing three different leptonic tensors:  $\tilde{L}_{\gamma\gamma}^{\mu\nu}$ , which is the purely electromagnetic tensor familiar from equation (3),  $\tilde{L}_{ZZ}^{\mu\nu}$  which describes the scattering via Z-boson exchange, and the two equivalent interference tensors  $\tilde{L}_{\gamma Z}^{\mu\nu} = \tilde{L}_{Z\gamma}^{\mu\nu}$ . The ZZ-case gives:

$$\begin{aligned} \tilde{L}_{ZZ}^{\mu\nu} &= 2 \sum_{\text{spins}} [\bar{u}(k') (C_L \gamma^\mu P_L + C_R \gamma^\mu P_R) u(k)] [\bar{u}(k') (C_L \gamma^\mu P_L + C_R \gamma^\mu P_R) u(k)] \\ &= 4(C_L^2 + C_R^2)(k^\mu k'^\nu + k'^\mu k^\nu - (k \cdot k') g^{\mu\nu}) + 4i(C_L^2 - C_R^2) \epsilon^{\mu\nu\rho\sigma} k_\rho k'_\sigma. \end{aligned} \quad (43)$$



The  $\gamma Z$  and the  $Z\gamma$  case each contribute in the form of:

$$\begin{aligned}\tilde{L}_{\gamma Z}^{\mu\nu} &= \sum_{\text{spins}} [\bar{u}(k')\gamma^\mu u(k)][\bar{u}(k')(C_L\gamma^\mu P_L + C_R\gamma^\mu P_R)u(k)] \\ &= 2(C_L + C_R)(k^\mu k'^\nu + k'^\mu k^\nu - (k \cdot k')g^{\mu\nu}) + 2i(C_L - C_R)\epsilon^{\mu\nu\rho\sigma}k_\rho k'_\sigma.\end{aligned}\quad (44)$$

Rewriting all leptonic tensors in terms of axial and vector couplings yields

$$\begin{aligned}\tilde{L}_{\gamma\gamma}^{\mu\nu} &= 2(k^\mu k'^\nu + k'^\mu k^\nu - (k \cdot k')g^{\mu\nu}) \\ \tilde{L}_{\gamma Z}^{\mu\nu} + \tilde{L}_{Z\gamma}^{\mu\nu} &= 4g_v(k^\mu k'^\nu + k'^\mu k^\nu - (k \cdot k')g^{\mu\nu}) + 4ia_v\epsilon^{\mu\nu\rho\sigma}k_\rho k'_\sigma \\ \tilde{L}_{ZZ}^{\mu\nu} &= 2(g_v^2 + a_v^2)(k^\mu k'^\nu + k'^\mu k^\nu - (k \cdot k')g^{\mu\nu}) + 4ia_v g_v\epsilon^{\mu\nu\rho\sigma}k_\rho k'_\sigma.\end{aligned}\quad (45)$$

Plugging this into (41) we get:

$$\begin{aligned}\frac{d\sigma}{dE'd\Omega} &= \frac{\alpha^2}{q^4} \frac{E'}{E} \left[ \left\{ \eta_{\gamma\gamma} W_{\gamma\gamma}^{\mu\nu} + 2g_v\eta_{\gamma Z} W_{\gamma Z}^{\mu\nu} + (g_v^2 + a_v^2)\eta_{ZZ} W_{ZZ}^{\mu\nu} \right\} \cdot 2(k^\mu k'^\nu + k'^\mu k^\nu - (k \cdot k')g^{\mu\nu}) \right. \\ &\quad \left. + 2i \left\{ 2a_v\eta_{\gamma Z} W_{\gamma Z}^{\mu\nu} + 2g_v a_v\eta_{ZZ} W_{ZZ}^{\mu\nu} \right\} \epsilon^{\mu\nu\rho\sigma} k_\rho k'_\sigma \right],\end{aligned}\quad (46)$$

where we can identify the content of the first curly bracket with the previously used hadronic tensor (26) and the content of the second curly bracket with the previously omitted antisymmetric part of the full tensor (24). Plugging in  $\frac{1}{2M^2}W_3\epsilon^{\mu\nu\kappa\lambda}p_\kappa q_\lambda$  for this bracket results in the Z-boson contribution

$$\begin{aligned}-\frac{W_3}{M^2}\epsilon^{\mu\nu\kappa\lambda}\epsilon_{\mu\nu\rho\sigma}p_\kappa q_\lambda k^\rho k'^\sigma &= 2\frac{W_3}{M^2}(\delta_\rho^\kappa\delta_\sigma^\lambda - \delta_\sigma^\kappa\delta_\rho^\lambda)p_\kappa q_\lambda k^\rho k'^\sigma \\ &= 2\frac{W_3}{M^2}k \cdot k'((p \cdot k) + (p \cdot k')) \\ &= 4\frac{W_3}{M}EE'(E + E')\sin^2\frac{\theta}{2}.\end{aligned}\quad (47)$$

By using (47) and the result of chapter 2.3 for the  $(ep \rightarrow eX)$ -cross section (37) for the first part in the square brackets of (46) we get

$$\frac{d\sigma}{dE'd\Omega} = 4\frac{\alpha^2}{q^4}E'^2 \left[ 2W_1 \sin^2\frac{\theta}{2} + W_2 \cos^2\frac{\theta}{2} + \frac{(E + E')}{M}W_3 \sin^2\frac{\theta}{2} \right]. \quad (48)$$

This is commonly rewritten in terms of dimensionless structure functions  $F_1, F_2, F_3$ :

$$F_1 = MW_1, \quad F_2 = \frac{q \cdot p}{M}W_2, \quad F_3 = \frac{q \cdot p}{M}W_3. \quad (49)$$

Using these structure functions, the differential cross section reads

$$\frac{d\sigma}{dE'd\Omega} = 4\frac{\alpha^2}{q^4}E'^2 \left[ 2\frac{F_1}{M} \sin^2\frac{\theta}{2} + \frac{F_2}{E - E'} \cos^2\frac{\theta}{2} + \frac{(E + E')}{M(E - E')}F_3 \sin^2\frac{\theta}{2} \right]. \quad (50)$$

Another important way of expressing this cross section is differential in  $x$  and  $Q^2$  instead of  $E'$

and  $\Omega$ . This can be achieved by using

$$\frac{d\sigma}{dE'd\Omega} = \frac{4\pi x E'}{y} \frac{d\sigma}{dx dQ^2}, \quad xy = \frac{2E'}{M} \sin^2 \frac{\theta}{2}, \quad 1 - y = \frac{E'}{E}. \quad (51)$$

The result is the form:

$$\frac{d\sigma}{dx dQ^2} = \frac{4\pi\alpha^2}{xQ^2} \left( F_1 xy^2 + F_2(1 - y - \frac{Mxy}{2E}) \pm y(1 - \frac{y}{2})x F_3 \right), \quad (52)$$

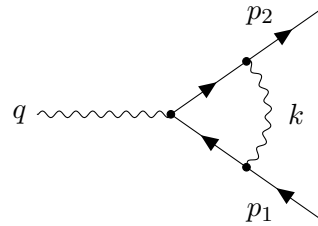
where the  $(-)$  case of the last term is for the case of an electron scattering on an anti-quark and originates from the parity violation of weak interactions. This distinction makes the structure function  $F_3$  especially useful for the determination of sea quark distributions. Lastly the cross section can also be expressed in terms of the longitudinal structure function  $F_L = F_2 - 2xF_1$ :

$$\frac{d\sigma}{dx dQ^2} = \frac{4\pi\alpha^2}{xQ^2} \left( -y^2 F_L + Y_+ F_2 \pm Y_- x F_3 \right), \quad (53)$$

where  $Y_{\pm} = 1 \pm (1 - y)^2$ . This structure function  $F_L$  can be used to test perturbative QCD because it is highly sensitive to the proton's gluon distribution at low  $x$  as discussed for example in [24].

## 2.5 QED vertex correction

Other corrections to the scattering process come from higher order calculations that can be computed from loop diagrams. Such calculations commonly require regularization as described in A.1.1. As an example of such a correction we will now calculate the vertex correction in massless QED at one-loop order because the results can be reused later on in section 2.7. There is only one contributing diagram:


(54)

with the invariant amplitude  $\mathcal{M}$  in  $D$  dimensions

$$\begin{aligned} \mathcal{M} &= -ie\mu^{\frac{4-D}{2}} \bar{u}(p_1) \Gamma^\mu v(p_2) \\ &= -\left(e\mu^{\frac{4-D}{2}}\right)^3 \int \frac{d^D k}{(2\pi)^D} \frac{\bar{u}(p_1) \gamma^\nu (\not{k} + \not{p}_1) \gamma^\mu (\not{k} - \not{p}_2) \gamma_\nu v(p_2)}{(k^2 + i\epsilon)((k + p_1)^2 + i\epsilon)((k - p_2)^2 + i\epsilon)}. \end{aligned} \quad (55)$$

Using the Feynman parameterization (A.103) for the denominator and rewriting the nominator we can write  $\Gamma^\mu$  as

$$\Gamma^\mu = -ie^2 \mu^{4-D} \int_0^1 dx \int_0^{1-x} dy \int \frac{d^D k}{(2\pi)^D} \frac{N^\mu(k, p_1, p_2)}{((k + xp_1 - yp_2)^2 + Q^2 xy + i\epsilon)^3}, \quad (56)$$

with

$$N^\mu = 2[(D-2)k^2 + 4k \cdot p_1 - 4k \cdot p_2 - 2Q^2]\gamma^\mu - 4[(D-2)k^\mu + 2p_1^\mu - 2p_2^\mu]k, \quad (57)$$

$$2p_1p_2 = (p_1 + p_2)^2 = q^2 = Q^2. \quad (58)$$

Then we can shift the integration variable  $k \rightarrow k - xp_1 + yp_2$  and get

$$\begin{aligned} \Gamma^\mu = & -2ie^2\mu^{4-D} \int_0^1 dx \int_0^{1-x} dy \int \frac{d^D k}{(2\pi)^D} \frac{1}{(k^2 + Q^2xy + i\epsilon)^3} \times \\ & \left[ \left( (D-2)(k^2 + (xp_1 - yp_2)^2) + 4yp_1p_2 - 4xp_1^2 - 4yp_2^2 + 4xp_1p_2 - 2Q^2 \right) \gamma^\mu \right. \\ & \left. - 2((D-2)(k^\mu - yp_2^\mu + xp_1^\mu) + 2p_1^\mu - 2p_2^\mu) (k + xp_1 - yp_2) \right], \end{aligned} \quad (59)$$

where we have already dropped any terms in the nominator proportional to uneven powers of  $k$  due to the integral's symmetry. When we look at  $\bar{u}(p_1)\Gamma^\mu v(p_2)$  we can also drop any terms including  $\not{p}_1$  or  $\not{p}_2$ , because  $\bar{u}(p_1)\not{p}_1 = \not{p}_2 v(p_2) = 0$  in massless QED. This gives us

$$\begin{aligned} \bar{u}(p_1)\Gamma^\mu v(p_2) = & -2ie^2\mu^{4-D} \int_0^1 dx \int_0^{1-x} dy \int \frac{d^D k}{(2\pi)^D} \frac{1}{(k^2 + Q^2xy + i\epsilon)^3} \times \\ & \bar{u}(p_1)\gamma^\mu \left[ \frac{(D-2)^2}{D} k^2 - 2Q^2 \left( (D-2)\frac{xy}{2} - x - y + 1 \right) \right] v(p_2), \end{aligned} \quad (60)$$

where we have used

$$\int d^D k \frac{\not{k}k^\nu}{(k^2 + Q^2xy + i\epsilon)} = \frac{1}{D}\gamma^\mu \int d^D k \frac{k^2}{(k^2 + Q^2xy + i\epsilon)}. \quad (61)$$

Next we will only consider the part of the integral with the numerator proportional to  $k^2$ :

$$\Gamma_1^\mu \equiv -2ie^2\mu^{4-D}\gamma^\mu \int_0^1 dx \int_0^{1-x} dy \int \frac{d^D k}{(2\pi)^D} \frac{\frac{(D-2)^2}{2}k^2}{(k^2 + Q^2xy + i\epsilon)^3}. \quad (62)$$

The  $k$  integration can be done using (A.24). Plugging in  $D = 4 - 2\epsilon$  then gives

$$\Gamma_1^\mu = \frac{2e^2}{(4\pi)^2} \left( \frac{\mu^2}{-Q^2} \right)^\epsilon \gamma^\mu \frac{(1-\epsilon)^2}{\epsilon} \Gamma(1+\epsilon)(4\pi)^\epsilon \int_0^1 dx \int_0^{1-x} dy \left( \frac{1}{xy} \right)^\epsilon. \quad (63)$$

If we substitute  $y = (1-x)y'$  and evaluate the  $y$ -integral then the  $x$ -integration gives a  $B$ -function and we can use (A.106) to get

$$\begin{aligned} \Gamma_1^\mu = & \frac{2e^2}{(4\pi)^2} \left( \frac{\mu^2}{-Q^2} \right)^\epsilon \gamma^\mu \frac{(1-\epsilon)^2}{\epsilon} \Gamma(1+\epsilon)(4\pi)^\epsilon \frac{\Gamma(1-\epsilon)\Gamma(2-\epsilon)}{(1-\epsilon)\Gamma(3-2\epsilon)} \\ = & \frac{e^2}{(4\pi)^2} \left( \frac{4\pi\mu^2}{-Q^2} \right)^\epsilon \gamma^\mu \Gamma(1+\epsilon) \left[ \frac{1}{\epsilon} + 1 \right] + \mathcal{O}(\epsilon). \end{aligned} \quad (64)$$

It is important to note here that the divergence is of ultraviolet origin as the integral converges for any  $\epsilon > 0$ .

The remaining part of the integral (60)

$$\Gamma_2^\mu = 2ie^2\mu^{4-D}\gamma^\mu \int_0^1 dx \int_0^{1-x} dy \int \frac{d^D k}{(2\pi)^D} \frac{2Q^2 \left((D-2)\frac{xy}{2} - x - y + 1\right)}{(k^2 + Q^2xy + i\epsilon)^3} \quad (65)$$

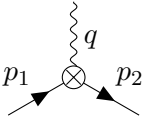
can be calculated in a similar manner by using (A.23) which gives

$$\Gamma_2^\mu = \frac{2e^2}{(4\pi)^2} \left( \frac{\mu^2}{-Q^2} \right)^\epsilon \gamma^\mu \Gamma(1+\epsilon)(4\pi)^\epsilon \int_0^1 dx \int_0^{1-x} dy (x+y-xy(1-\epsilon)-1) \left( \frac{1}{xy} \right)^{-1-\epsilon}. \quad (66)$$

Using the substitution  $y = (1-x)y'$  and the  $B$ -function like before results in

$$\begin{aligned} \Gamma_2^\mu &= \frac{2e^2}{(4\pi)^2} \left( \frac{4\pi\mu^2}{-Q^2} \right)^\epsilon \gamma^\mu \Gamma(1+\epsilon) \left[ \left( \frac{1}{\epsilon} + \frac{1}{1-\epsilon} \right) \frac{\Gamma(-\epsilon)\Gamma(2-\epsilon)}{\Gamma(2-2\epsilon)} - \frac{\Gamma(1-\epsilon)\Gamma(2-\epsilon)}{\Gamma(3-2\epsilon)} \right] \\ &= \frac{e^2}{(4\pi)^2} \left( \frac{4\pi\mu^2}{-Q^2} \right)^\epsilon \gamma^\mu \Gamma(1+\epsilon) \left[ -\frac{2}{\epsilon^2} - \frac{4}{\epsilon} - 9 + \frac{\pi^2}{3} \right]. \end{aligned} \quad (67)$$

In this case the divergences are all of infrared origin because the integral would converge for any  $\epsilon < 0$ . Now we want to use (on shell) renormalization of the electric charge to get rid of the ultraviolet divergence. The counterterm can be calculated from the diagram



$$= -i\delta_1 \left( e_R \mu^{\frac{4-D}{2}} \right) \bar{u}(p_2) \gamma^\mu v(p_1). \quad (68)$$

To obtain the counterterm we need to compute the divergent part of  $\Gamma^\mu$  at  $Q = 0$ . From (60) we then get:

$$\Gamma^\mu(Q^2 = 0) = -2ie_R^2\mu^{4-D}\gamma^\mu \frac{(D-2)^2}{D} \int \frac{d^D k}{(2\pi)^D} \frac{1}{k^4} \int_0^1 dx \int_0^{1-x} dy.$$

This integral is both IR and UV divergent. To evaluate it, we first need to introduce an arbitrary scale  $\Lambda$  to split the integral into UV and IR divergent regions:

$$\Gamma^\mu(Q^2 = 0) = -ie_R^2\mu^{4-D}\gamma^\mu \frac{(D-2)^2}{D} \frac{i\Omega_D}{(2\pi)^D} \left( \int_0^\Lambda dk k^{D-5} + \int_\Lambda^\infty dk k^{D-5} \right)$$

We can then take  $D = 4 - 2\epsilon_{\text{IR}}$  with  $\epsilon_{\text{IR}} < 0$  for the first integral and  $D = 4 - 2\epsilon_{\text{UV}}$  with  $\epsilon_{\text{UV}} > 0$  for the second integral:

$$\begin{aligned}\Gamma^\mu(Q^2 = 0) &= e_R^2 \mu^{4-D} \gamma^\mu \frac{(D-2)^2}{D} \frac{\Omega_D}{(2\pi)^D} \left( \ln \Lambda - \frac{1}{2\epsilon_{\text{IR}}} + \frac{1}{2\epsilon_{\text{UV}}} - \ln \Lambda \right) \\ &= e_R^2 \mu^{4-D} \gamma^\mu \frac{(D-2)^2}{D} \frac{\pi^{D/2}}{\Gamma(\frac{D}{2})(2\pi)^D} \left( -\frac{1}{\epsilon_{\text{IR}}} + \frac{1}{\epsilon_{\text{UV}}} \right) \\ &= \frac{e_R^2}{(4\pi)^2} \mu^{4-D} \gamma^\mu \left( \frac{1}{\epsilon_{\text{UV}}} - \frac{1}{\epsilon_{\text{IR}}} \right) + \mathcal{O}(\epsilon^0)\end{aligned}\tag{69}$$

$$\implies \delta_1 = -\frac{e_R^2}{(4\pi)^2} \mu^{4-D} \left( \frac{1}{\epsilon_{\text{UV}}} - \frac{1}{\epsilon_{\text{IR}}} \right)\tag{70}$$

The renormalization cancels the ultraviolet divergence of  $\Gamma_1^\mu$  without affecting the finite parts in any way. Combining  $\Gamma_1^\mu$ ,  $\Gamma_2^\mu$  and the counterterm now leaves us with a result that is entirely infrared divergent:

$$\Gamma_R^\mu = \Gamma_1^\mu + \Gamma_2^\mu + \delta_1 \gamma^\mu \equiv \gamma^\mu f(Q^2)\tag{71}$$

with

$$f(Q^2) = \frac{e_R^2}{(4\pi)^2} \left( \frac{4\pi\mu^2}{-Q^2} \right)^\epsilon \Gamma(1+\epsilon) \left( -\frac{2}{\epsilon^2} - \frac{3}{\epsilon} - 8 + \frac{\pi^2}{3} \right).\tag{72}$$

This result will become very important later on, when deriving the DGLAP evolution equations, where we will use it in the form

$$f(Q^2) = \frac{\alpha}{4\pi} \left( \frac{4\pi\mu^2}{-Q^2} \right)^\epsilon \frac{\Gamma(1-\epsilon)}{\Gamma(1-2\epsilon)} \left( -\frac{2}{\epsilon^2} - \frac{3}{\epsilon} - 8 - \frac{\pi^2}{3} \right).\tag{73}$$

where we have neglected any terms of order  $\epsilon$  or higher.

## 2.6 The parton model

When performing scattering experiments at increasingly high energies in the 1960s it was observed that the high energy limit of the cross section (37) has the form (14) which was derived assuming scattering on pointlike fermions. This observation was what lead to the invention of the parton model which postulates that the proton is built up from pointlike constituents called partons. Subsequently these partons were identified with the quarks of the quark model which are held together by gluons.

### 2.6.1 Basic statements and consequences of the parton model

The core statement of the parton model is that for small enough wavelengths the electron no longer scatters on the proton as a whole but instead performs an elastic scattering on a "free"

quark within the proton. Therefore we can see from (35) and (37) that the proton structure functions become

$$2W_1 = \frac{Q^2}{2m^2} \delta\left(\nu - \frac{Q^2}{2m}\right) \quad \text{and} \quad W_2 = \delta\left(\nu - \frac{Q^2}{2m}\right), \quad (74)$$

where  $m$  is the mass of the quark. Using  $\delta(\frac{a}{b}) = |b|\delta(a)$ ,  $W_1$  and  $W_2$  can be rewritten so that they only depend on the dimensionless ratio  $\frac{Q^2}{2m\nu}$  but not on  $Q^2$  and  $\nu$  separately:

$$\begin{aligned} 2mW_1(\nu, Q^2) &= \frac{Q^2}{2m\nu} \delta\left(1 - \frac{Q^2}{2m\nu}\right), \\ \nu W_2(\nu, Q^2) &= \delta\left(1 - \frac{Q^2}{2m\nu}\right). \end{aligned} \quad (75)$$

This behaviour is called Bjorken scaling<sup>3</sup> [25], [26] and cannot be replicated by the formula for elastic  $ep$ -scattering (21). For simplicity we just look at the case of  $\kappa = 0$ , which means  $G_E = G_M \equiv G$ . In this case comparing (36) and (37) gives:

$$\begin{aligned} W_1^{\text{elastic}} &= \frac{Q^2}{4M^2} G^2(Q^2) \delta\left(\nu - \frac{Q^2}{2M}\right), \\ W_2^{\text{elastic}} &= G^2(Q^2) \delta\left(\nu - \frac{Q^2}{2M}\right). \end{aligned} \quad (76)$$

The  $Q^2$ -dependence of  $G(Q^2)$  makes it impossible to write the structure functions as functions of a single dimensionless variable. In the elastic case there is also an explicit mass scale, set by the empirical value for the inverse size of the proton in the dipole formula (22) for  $G(Q^2)$ . When  $Q^2$  goes to higher values we can see in figure 2.1 that  $G(Q^2)$  represses the chance of elastic scattering and the proton becomes more likely to break up.

By introducing the new dimensionless variable  $\omega = \frac{2q \cdot p}{Q^2} = \frac{2M\nu}{Q^2}$ , we can summarize the behavior for large  $Q^2$  as:

$$\begin{aligned} MW_1(\nu, Q^2) &\xrightarrow{\text{large } Q^2} F_1(\omega), \\ \nu W_2(\nu, Q^2) &\xrightarrow{\text{large } Q^2} F_2(\omega), \end{aligned} \quad (77)$$

Note that we used the proton mass  $M$  instead of the quark mass  $m$  to define  $\omega$  because it will be more convenient later on. The existence of point-like, spin- $\frac{1}{2}$  particles inside the proton can therefore be observed experimentally by looking at either structure function  $W_1$  or  $W_2$  for a fixed  $\omega$ , while varying  $Q^2$ . Figure 2.4 shows an example of data from the BCDMS experiment [20] showing no signs of  $Q^2$  dependence. As we will see later on in chapter 2.7, specifically figure 2.9 this Bjorken scaling is actually violated and there is a logarithmic  $Q^2$  dependence arising from higher order QCD corrections. For now we will ignore this however and look only at the leading order parton model.

---

<sup>3</sup>Named after James Bjorken, who was the first to predict this behavior in the late 1960s.

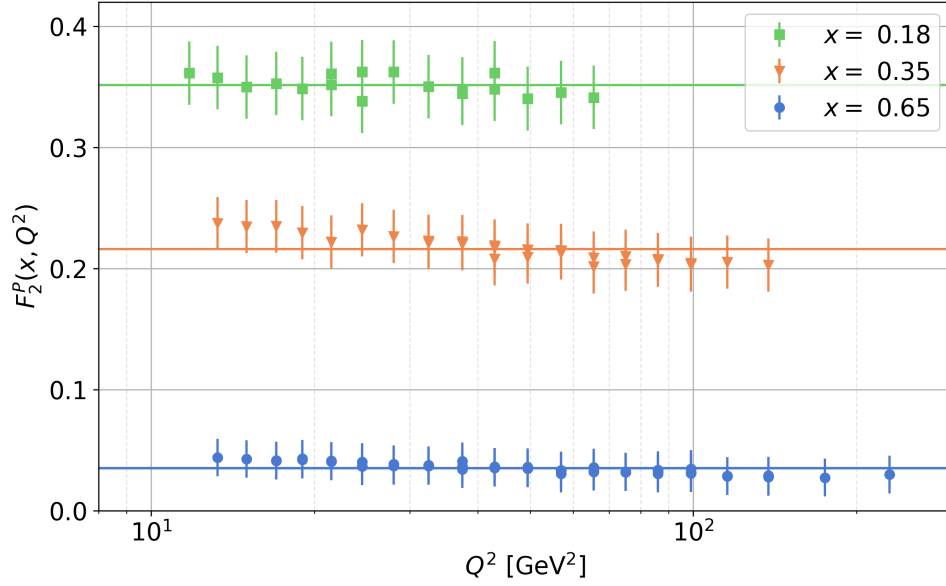


Figure 2.4:  $F_2$  as a function of  $Q^2$  for three different values of the Bjorken variable  $x$ .

Since there are quarks with different charges and also gluons inside the proton the interaction can be pictured as

$$E, p \Rightarrow \left[ \text{Diagram: Photon hitting a proton vertex} \right] = \sum_i \int dx e_i^2 \left[ \text{Diagram: Photon hitting a parton } i \text{ inside a proton} \right], \quad (78)$$

where  $i \in \{\text{u-quark, d-quark, ..., gluon}\}$  and the respective charges are given by  $e_i$ . The gluons have an electric charge of 0 and can therefore be ignored for now. We can then introduce the parton momentum distribution  $f_i(x)$  which describes the probability that the parton  $i$  that interacts with the photon carries a fraction  $x$  of the protons total momentum  $p$ . Denoting the total chance of scattering on a parton  $i$  as  $P_i$ , we can define

$$f_i(x) \equiv \frac{dP_i}{dx} = \left. p \Rightarrow \left[ \text{Diagram: Parton } i \text{ carrying momentum } xp \text{ and the rest of the proton carrying } (1-x)p \right] \right\}.$$

The integral over the momentum fractions weighted by their probabilities has to be 1 due to conservation of momentum:

$$\sum_i \int_0^1 dx x f_i(x) = 1. \quad (79)$$

The kinematics of the parton are:

Table 1: Parton kinematics<sup>4</sup>.

	<b>Proton</b>	<b>Parton</b>
Energy	$E$	$xE$
Momentum	$p$	$xp$
Mass	$M$	$m = (x^2 E^2 - x^2 p^2)^{\frac{1}{2}} = xM$

Using this, along with (75) and (77) we get:

$$\begin{aligned} \frac{2F_1(\omega)}{M} = \frac{Q^2}{2m^2\nu} \delta\left(1 - \frac{Q^2}{2m\nu}\right) &\implies F_1(\omega) = \frac{Q^2}{4xm\nu} \left(1 - \frac{Q^2}{2m\nu}\right) = \frac{1}{2x\omega} \delta\left(x - \frac{1}{\omega}\right), \\ F_2(\omega) = \delta\left(1 - \frac{Q^2}{2m\nu}\right) &\implies F_2(\omega) = x\delta\left(x - \frac{1}{\omega}\right), \end{aligned} \quad (80)$$

for an electron scattering on a parton with momentum fraction  $x$  and unit charge. These results can then be summed over all partons and integrated over  $x$  to obtain the proton's structure functions

$$\begin{aligned} F_1(\omega) &= \sum_i \int dx \, e_i^2 \frac{1}{2\omega x} f_i(x) \delta\left(x - \frac{1}{\omega}\right), \\ F_2(\omega) &= \sum_i \int dx \, e_i^2 x f_i(x) \delta\left(x - \frac{1}{\omega}\right). \end{aligned} \quad (81)$$

Conventionally  $F_{1,2}(\omega)$  are expressed as  $F_{1,2}(x)$ . Looking back at (77) we see that at large  $Q^2$  the protons structure functions become

$$\begin{aligned} MW_1(\nu, Q^2) &\xrightarrow{\text{large } Q^2} F_1(x) = \frac{1}{2x} \sum_i e_i^2 x f_i(x), \\ \nu W_2(\nu, Q^2) &\xrightarrow{\text{large } Q^2} F_2(x) = \sum_i e_i^2 x f_i(x), \end{aligned} \quad (82)$$

with

$$x = \frac{1}{\omega} = \frac{Q^2}{2M\nu}. \quad (83)$$

The structure functions of the proton depend only on the dimensionless variable  $x$ , often called Bjorken  $x$ . This independence of the experimental energy scale implies that the partons are point-like because higher energy allows better spatial resolution and the only structure that is invariant under spatial scaling is a point. The equations (82) also show us the Callan-Gross relation  $F_1(x) = \frac{1}{2x} F_2(x)$ .

<sup>4</sup>The mass of the partons cannot actually be variable like  $m = xM$  suggests. Assuming the parton's momentum is  $xp$ , then the energy can only be  $xE$  if the masses are  $m = M = 0$ . We can justify this, however, by working in a frame of reference where  $|p| \gg m, M$ . In a frame where the protons momentum is infinite, the kinematics become exact.



### 2.6.2 The quark structure of nucleons

The results from the previous section allow us to gain insight into the quark structure of nucleons. Disregarding charm-, bottom-, and top-quarks because of their high mass we can write  $F_2(x)$  from (82) as

$$\begin{aligned}\frac{1}{x}F_2^{ep}(x) &= \left(\frac{2}{3}\right)^2 [u^p(x) + \bar{u}^p(x)] + \left(\frac{1}{3}\right)^2 [d^p(x) + \bar{d}^p(x)] + \left(\frac{1}{3}\right)^2 [s^p(x) + \bar{s}^p(x)], \\ \frac{1}{x}F_2^{en}(x) &= \left(\frac{2}{3}\right)^2 [u^n(x) + \bar{u}^n(x)] + \left(\frac{1}{3}\right)^2 [d^n(x) + \bar{d}^n(x)] + \left(\frac{1}{3}\right)^2 [s^n(x) + \bar{s}^n(x)],\end{aligned}\tag{84}$$

for protons and neutrons respectively. Since the proton and neutron form an isospin doublet their quark contents are related by

$$\begin{aligned}u^p(x) &= d^n(x) \equiv u(x), \\ d^p(x) &= u^n(x) \equiv d(x), \\ s^p(x) &= s^n(x) \equiv s(x).\end{aligned}\tag{85}$$

Further, we can divide the quarks into valence quarks  $q_v$  and sea quarks  $q_s$ . Sea quarks appear inside any hadrons as quark-antiquark pairs generated from gluons. Since they always appear in  $q\bar{q}$ -pairs they do not affect the total quantum numbers of the nucleons. Assuming equal mass for up-, down- and strange-quark it is convenient to write

$$\begin{aligned}u_s(x) &= \bar{u}_s(x) = d_s(x) = \bar{d}_s(x) = s_s(x) = \bar{s}_s(x) = S(x), \\ u(x) &= u_v(x) + u_s(x), \\ d(x) &= d_v(x) + d_s(x),\end{aligned}\tag{86}$$

so that we can bring  $F_2^{ep}$  and  $F_2^{en}$  into the form

$$\frac{1}{x}F_2^{ep} = \frac{1}{9}(4u_v + d_v) + \frac{4}{3}S \quad \text{and} \quad \frac{1}{x}F_2^{en} = \frac{1}{9}(u_v + 4d_v) + \frac{4}{3}S.\tag{87}$$

Summing over the contributions of all partons the proton's quantum numbers must be recovered:

$$\int_0^1 dx x [u(x) - \bar{u}(x)] = 2,\tag{88}$$

$$\int_0^1 dx x [d(x) - \bar{d}(x)] = 1,\tag{89}$$

$$\int_0^1 dx x [s(x) - \bar{s}(x)] = 0.\tag{90}$$

This is the only possible combination yielding the correct charge ( $2 \cdot \frac{2}{3} - 1 \cdot \frac{1}{3} = 1$ ), baryon number ( $3 \cdot \frac{1}{3} = 1$ ) and strangeness 0. Since the sea quarks are created by gluons their number is expected to grow logarithmically as  $x \rightarrow 0$ . Therefore we expect the three valence quarks to be overshadowed by the multitude of sea quarks when looking at small momentum fractions  $x$ .

This means that

$$\frac{F_2^{en}(x)}{F_2^{ep}(x)} \xrightarrow{x \rightarrow 0} 1. \quad (91)$$

In the opposite case of large momentum fractions, it is very unlikely for a sea pair to be created with a large momentum fraction and therefore the probability of interacting with valence quarks heavily outweighs that of the sea quarks:

$$\frac{F_2^{en}(x)}{F_2^{ep}(x)} \xrightarrow{x \rightarrow 1} \frac{u_v + 4d_v}{4u_v + d_v}. \quad (92)$$

The data shown in figure 2.5 confirms our prediction from (91) and with (92) it hints that  $u_v \gg d_v$  at large  $x$  as the ratio (92) tends towards  $\frac{1}{4}$ . Note that this means that  $u_v(x) \neq 2d_v(x)$  and therefore they have to look qualitatively different to satisfy the sum rules (88) - (90).

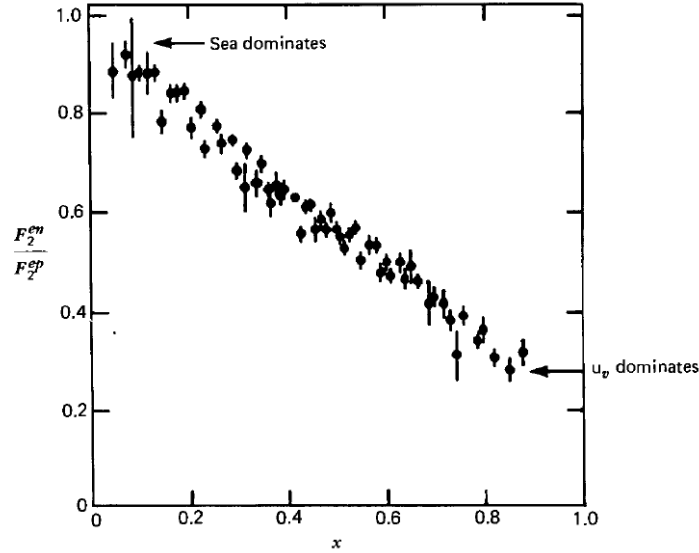


Figure 2.5: The ratio  $F_2^n/F_2^p$  as a function of  $x$ . Data are from the Stanford Linear Accelerator, figure from [23].

Figure 2.6 shows a qualitative estimate of  $F_2$  for various cases of proton structure. Non-interacting quarks would obviously lead to sharp peaks at  $x = \frac{1}{N_{\text{quarks}}}$ , while interaction through gluons smears out that peak towards a continuous distribution vanishing at both  $x = 0$  and  $x = 1$ . If one also considers the splitting of gluons into sea quarks the small  $x$  region becomes non-vanishing. We can remove the sea quark part from data by looking at the difference between the proton and neutron structure functions  $F_2^{ep} - F_2^{en}$  so that the result is independent of the sea  $S(x)$  according to (87). Looking at the data in figure 2.7 this does indeed look like the prediction from 2.6 for the case without sea quarks.

Alternatively,  $F_2^{ep,en}$  data can be used to fit the quark structure functions, as is the main goal of this thesis. An example of this is shown in figure 2.8. On the left side of 2.8 we can clearly see how  $u(x) = u_v(x) + u_s(x)$  approaches  $u_s(x)$  as  $x \rightarrow 0$  and that the qualitative shape of  $F_2$  from the last prediction of figure 2.6 would be replicated by these quark distributions.

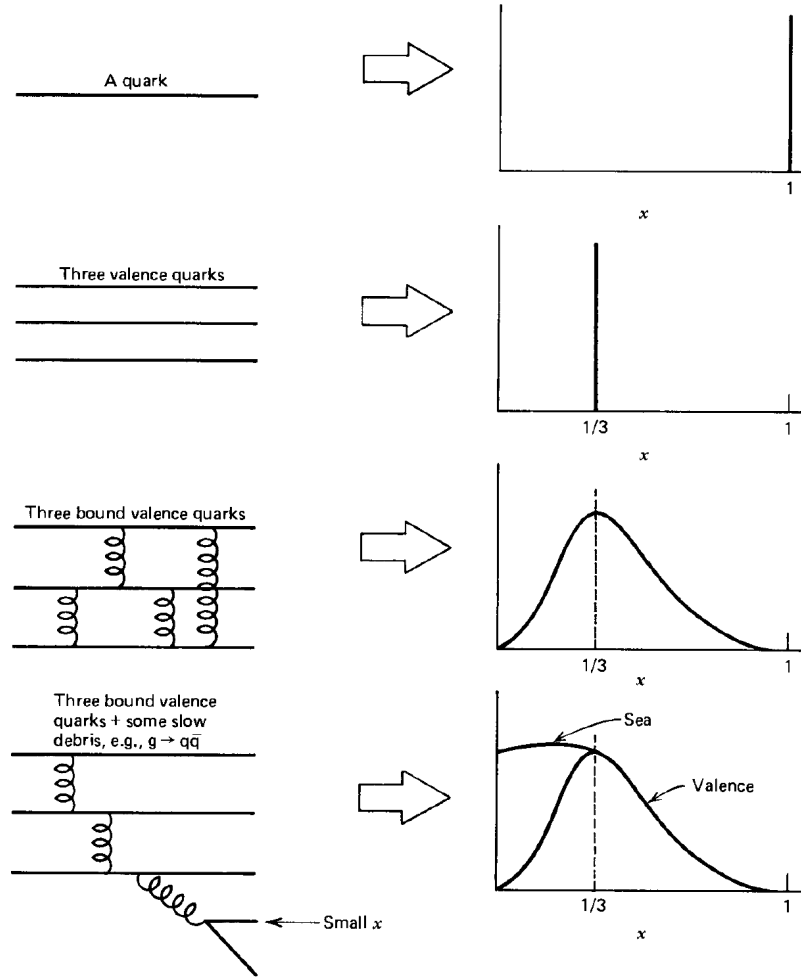


Figure 2.6: Qualitative expectation for the structure function  $F_2^p$  based on different assumed proton compositions. Figure taken from [23].

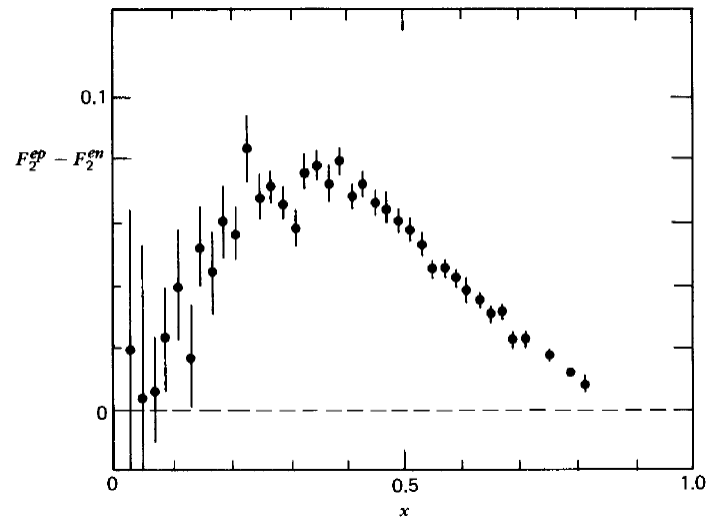


Figure 2.7: The difference  $F_2^n - F_2^p$  as a function of  $x$ . Data are from the Stanford Linear Accelerator, figure from [23].

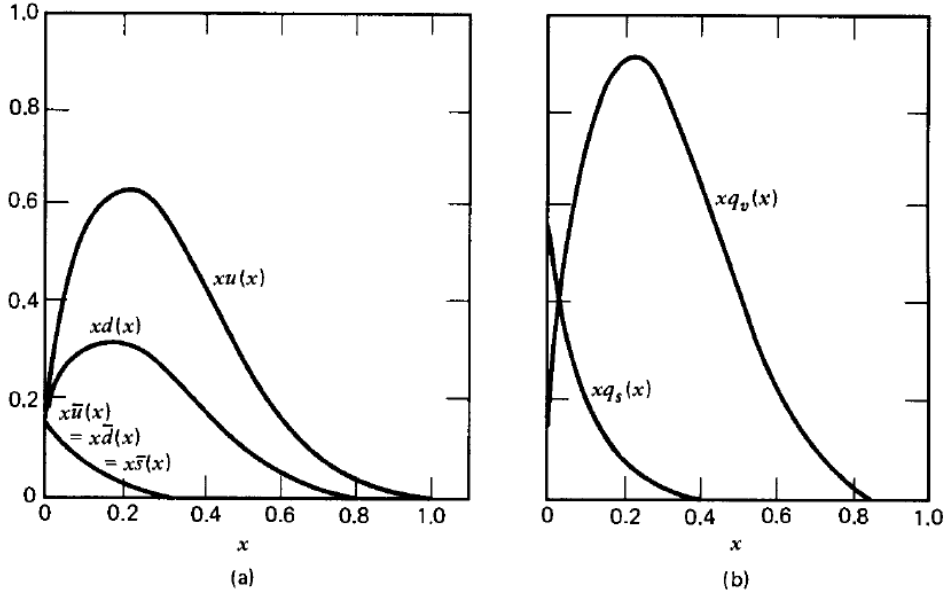


Figure 2.8: Quark structure functions fitted to  $F_2$  data from the Stanford Linear Accelerator. The right plot shows the total valence and sea quark contributions. Figure taken from [23]

From the data in figure 2.8 we can also gain some insight about the gluons, even though their neutral electric charge causes them to have no impact on the structure functions  $F_2$ . Integrating the experimental data over  $x$  (while neglecting the small strange-quark contributions) we get

$$\begin{aligned} \int dx F_2^{ep}(x) &= \frac{4}{9}\epsilon_u(u + \bar{u}) + \frac{1}{9}\epsilon_d = 0.18, \\ \int dx F_2^{en}(x) &= \frac{1}{9}\epsilon_u(u + \bar{u}) + \frac{4}{9}\epsilon_d = 0.12, \end{aligned} \quad (93)$$

where

$$\epsilon_q \equiv \int_0^1 dx x(q + \bar{q}) \quad (94)$$

is the total momentum fraction carried by quarks and antiquarks of a specific flavor. Because the sum of the momentum fractions of all partons has to be 1 we can estimate the momentum fraction carried by gluons to be

$$\epsilon_g \approx 1 - (\epsilon_u + \epsilon_d), \quad (95)$$

which results in

$$\epsilon_u = 0.36, \quad \epsilon_d = 0.18, \quad \epsilon_g = 0.46. \quad (96)$$

Perhaps unexpectedly a substantial portion of the protons total momentum is carried by the massless gauge bosons mediating the strong force between the quarks.

## 2.7 The QCD improved parton model

The naive parton model from earlier has predicted Bjorken scaling and the Callan-Gross relation, which were both observed experimentally. Newer data like 2.9 shows however, that there actually is a  $Q^2$  dependence in the structure functions. This logarithmic dependence can be calculated in perturbative QCD.

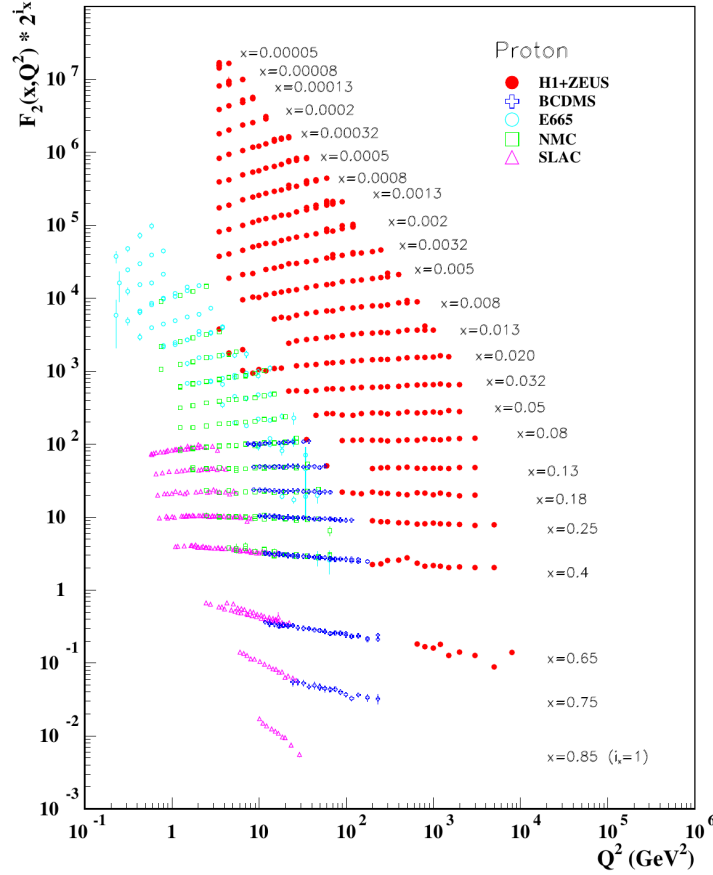


Figure 2.9:  $F_2^p$  measured in various experiments. The values of  $F_2^p$  for the various  $x$  are offset from each other by multiplication with  $2^i$ , where  $i$  counts the  $x$  bins in descending order. Figure taken from [27].

To explain this, we will first write down  $\hat{W}^{\mu\nu}(z, Q^2)$  as the partonic version of  $W^{\mu\nu}(x, Q^2)$ , where  $z \equiv \frac{Q^2}{2p_i \cdot q}$ . We assume that the naive parton model holds insofar that the probability of finding  $p_i^\mu = \xi p^\mu$  for any  $0 \leq \xi \leq 1$  is given by a parton distribution function  $f_i(\xi)$ , so that  $x = z\xi$  and we integrate over  $\xi$ :

$$\begin{aligned} W^{\mu\nu}(x, Q^2) &= \sum_i \int_0^1 dz \int_0^1 d\xi f_i(\xi) \hat{W}^{\mu\nu}(z, Q^2) \delta(x - z\xi) \\ &= \sum_i \int_x^1 \frac{d\xi}{\xi} f_i(\xi) \hat{W}^{\mu\nu}\left(\frac{x}{\xi}, Q^2\right). \end{aligned} \quad (97)$$

In leading-order QCD the partonic tensor  $\hat{W}^{\mu\nu}(z, Q^2)$  is entirely characterized by the process

$\gamma^* q \rightarrow q$ :

$$\begin{aligned}\hat{W}^{\mu\nu}(z, Q^2) &= \frac{e_i^2}{2} \int \frac{d^3 \vec{p}_f}{(2\pi)^3} \frac{1}{2E_f} \text{Tr} [\gamma^\mu \not{p}_i \gamma^\nu \not{p}_f] (2\pi)^4 \delta^4(p_i + q - p_f) \\ &= 2\pi Q_i^2 \left[ \left( -g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2} \right) + \frac{4z^2}{Q^2} \left( p_i^\mu - \frac{p_i \cdot q}{q^2} q^\mu \right) \left( p_i^\nu - \frac{p_i \cdot q}{q^2} q^\nu \right) \right] \delta(1-z). \quad (98)\end{aligned}$$

Comparing this to the hadronic tensor from (26) we find

$$\hat{W}_1 = 2\pi e_i^2 \delta(1-z) \quad \text{and} \quad \hat{W}_2 = \frac{8\pi e_i^2}{Q^2} z^2 \delta(1-z). \quad (99)$$

This confirms that the Callan-Gross relation  $\hat{W}_1 = \frac{Q^2}{4z^2} \hat{W}_2$  still holds in leading-order QCD.

Now we introduce the form factor  $W_0 \equiv -g^{\mu\nu} W_{\mu\nu}$ , which is just the unpolarized cross section for the process  $\gamma^* p \rightarrow X$ . For the hadronic tensor this evaluates to

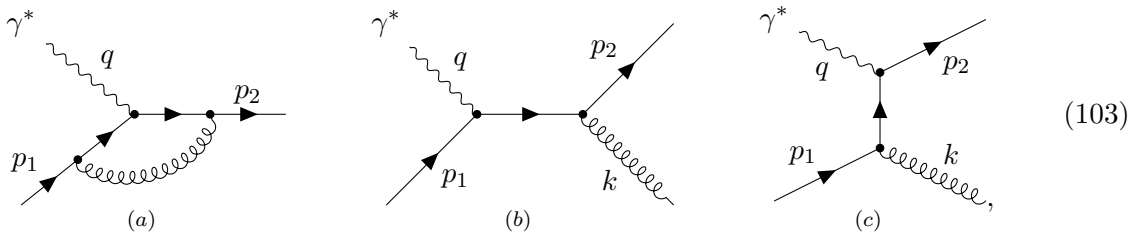
$$W_0(x, Q^2) = 3W_1(x, Q^2) - W_2(x, Q^2) \left( m_p^2 + \frac{Q^2}{4x^2} \right), \quad (100)$$

which can be further simplified in the case of  $Q \gg m_p$  to

$$W_0(x, Q^2) = 3W_1(x, Q^2) - \frac{Q^2}{4x^2} W_2(x, Q^2) = 2W_1(x, Q^2) \quad (101)$$

$$\longrightarrow W_0(x, Q^2) = 4\pi \sum_i e_i^2 f_i(x) = \frac{4\pi}{x} F_2(x, Q^2). \quad (102)$$

This equation can be used to define PDFs even beyond leading-order. This allows us to calculate the  $Q^2$  dependence of the PDFs a convenient way without the need to look at the individual structure functions. At leading order we know from (99) and (101) that  $W_0^{\text{LO}} = 4\pi e_i^2 \delta(1-z)$ . To calculate  $W_0$  at next-to-leading order (NLO), we need to consider the three contributing graphs:



where (a) is the virtual  $\gamma^* q \rightarrow q$  and (b) + (c) are real  $\gamma^* q \rightarrow qg$  graphs. We have already calculated the process  $\gamma^* \rightarrow q\bar{q}$  at 1-loop order in massless QED in chapter 2.5. To be able to use those results here we need to consider two things: What changes because of the rotation of the diagram and what changes when replacing the internal photon by a gluon. The rotation of the diagram does not change  $\Gamma(q^2, p_1^2, p_2^2)$ , but in the rotated diagram we have  $q^2 = -Q^2$  instead of  $q^2 = Q^2$ . To account for this we can simply replace  $f(Q^2)$  in (73) with  $f(-Q^2)$ . The second change requires a bit more effort because we have to compare the gluon interaction terms and

propagator with those of the photon:

$$\begin{array}{c} \text{photon vertex} \\ \text{---} \end{array} = -ie\gamma^\mu \longrightarrow \begin{array}{c} \text{gluon vertex} \\ \text{---} \end{array} = -ig\gamma^\mu T_{ij}^a \quad (104)$$

$$\text{photon propagator} = \frac{-ig_{\mu\nu}}{p^2 + i\epsilon} \longrightarrow \text{gluon propagator} = \frac{-ig_{\mu\nu}}{p^2 + i\epsilon} \delta^{ab} \quad (105)$$

$$\implies e^2 \longrightarrow \frac{1}{N_C} g^2 T_{ij}^a T_{ji}^a = g^2 C_F, \quad (106)$$

where the factor  $\frac{1}{N_C}$  comes from averaging over the initial state colors. This gives us in total:

$$f(-Q^2) = \frac{\alpha_s}{4\pi} C_F \left( \frac{4\pi\mu^2}{Q^2} \right)^\epsilon \frac{\Gamma(1-\epsilon)}{\Gamma(1-2\epsilon)} \left( -\frac{2}{\epsilon^2} - \frac{3}{\epsilon} - 8 - \frac{\pi^2}{3} \right), \quad (107)$$

which we can then use to calculate  $\hat{W}_0^V$ :

$$\hat{W}_0^{\text{LO}} + \hat{W}_0^V = (1 + f(-Q^2))^2 \hat{W}_0^{\text{LO}} = \hat{W}_0^{\text{LO}} + 2\hat{W}_0^{\text{LO}} f(-Q^2) + \text{higher order terms} \quad (108)$$

$$\implies \hat{W}_0^V = 4\pi e_i^2 \frac{\alpha_s}{2\pi} C_F \left( \frac{4\pi\mu^2}{Q^2} \right)^\epsilon \frac{\Gamma(1-\epsilon)}{\Gamma(1-2\epsilon)} \left( -\frac{2}{\epsilon^2} - \frac{3}{\epsilon} - 8 - \frac{\pi^2}{3} \right) \delta(1-z). \quad (109)$$

### 2.7.1 Real emission of gluons

Having calculated the virtual graph, we still need the two real emission graphs. In  $D$  dimensions the two contributing diagrams give:

$$\begin{array}{c} \text{Diagram 1: } \gamma^* \text{ splits into } q \text{ and } p_1, \text{ } q \text{ splits into } p_2 \text{ and } k, \text{ } p_2 \text{ splits into } i \text{ and } a. \\ \text{---} \end{array} = (-ie)(-igT_{ij}^a) \mu^{\frac{4-D}{2}} \epsilon_q^\mu \epsilon_k^{*\nu} \bar{u}(p_2) \gamma^\nu \frac{i(\not{q} + \not{p}_1 + m)}{(q+p_1)^2 - m^2} \gamma^\mu u(p_1) \quad (110)$$

and

$$\begin{array}{c} \text{Diagram 2: } \gamma^* \text{ splits into } q \text{ and } p_2, \text{ } q \text{ splits into } p_1 \text{ and } k, \text{ } p_1 \text{ splits into } j \text{ and } a. \\ \text{---} \end{array} = (-ie)(-igT_{ij}^a) \mu^{\frac{4-D}{2}} \epsilon_q^\mu \epsilon_k^{*\nu} \bar{u}(p_2) \gamma^\mu \frac{i(\not{p}_1 - \not{k} + m)}{(p_1 - k)^2 - m^2} \gamma^\nu u(p_1). \quad (111)$$

The total amplitude is the sum of the amplitudes for the s-channel  $M_s$  and t-channel  $M_t$ :

$$\begin{aligned}\mathcal{M} &= \mathcal{M}_s + \mathcal{M}_t \\ &= -egT_{ij}^a \mu^{\frac{4-D}{2}} \epsilon_q^\mu \epsilon_k^{*\nu} \bar{u}(p_2) \left[ \frac{\gamma^\nu (\not{q} + \not{p}_1 + m) \gamma^\mu}{s - m^2} + \frac{\gamma^\mu (\not{p}_1 - \not{k} + m) \gamma^\nu}{t - m^2} \right] u(p_1),\end{aligned}\quad (112)$$

where the Mandelstam variables are

$$s = q^2 + 2(p_1 \cdot q) = 2(p_2 \cdot k), \quad (113)$$

$$t = q^2 - 2(p_2 \cdot q) = 2(p_1 \cdot k), \quad (114)$$

$$u = q^2 - 2(q \cdot k) = 2(p_1 \cdot p_2), \quad (115)$$

$$\text{with } s + t + u = \sum_i m_i^2 = q^2. \quad (116)$$

We want to calculate the polarization summed  $|\mathcal{M}|^2$ . To do this we write  $\mathcal{M} = \epsilon_\mu M_\mu$  and use the photon polarization sum (A.126)

$$\sum_{\text{pols.}, i} |\mathcal{M}|^2 = -M_\mu^* M_\mu + \frac{1}{2E^2} (p_\mu M_\mu^* M_\nu \bar{p}_\nu + \bar{p}_\mu M_\mu^* M_\nu p_\nu). \quad (117)$$

Using the Ward identity  $p^\mu M_\mu = 0$  allows us to write

$$\sum_{\text{pols.}, i} \epsilon_\mu^{i*} \epsilon_\nu^i \longrightarrow -g_{\mu\nu}. \quad (118)$$

This only works with the sum of all relevant diagrams because the Ward identity relies on gauge invariance, which the individual diagrams do not necessarily fulfill. To simplify the calculations we go to the high energy limit where we can neglect the masses  $m$ . Using

$$\sum_{s=1}^2 v_s(p) \bar{v}_s(p) = \not{p} - m \quad \text{and} \quad \sum_{s=1}^2 u_s(p) \bar{u}_s(p) = \not{p} + m, \quad (119)$$

then gives us the four components

$$\sum_{\text{spins / pols.}} |\mathcal{M}_t|^2 = \frac{e^2 g^2 C_F N_C}{t^2} \mu^{4-D} \text{Tr} \left[ \not{p}_2 \gamma^\mu (\not{p}_1 - \not{k}) \gamma^\nu \not{p}_1 \gamma^\nu (\not{p}_1 - \not{k}) \gamma_\mu \right], \quad (120)$$

$$\sum_{\text{spins / pols.}} |\mathcal{M}_s|^2 = \frac{e^2 g^2 C_F N_C}{s^2} \mu^{4-D} \text{Tr} \left[ \not{p}_2 \gamma^\nu (\not{q} + \not{p}_1) \gamma^\mu \not{p}_1 \gamma^\mu (\not{q} + \not{p}_1) \gamma_\nu \right], \quad (121)$$

$$\sum_{\text{spins / pols.}} |\mathcal{M}_t^* \mathcal{M}_s| = \frac{e^2 g^2 C_F N_C}{st} \mu^{4-D} \text{Tr} \left[ \not{p}_2 \gamma^\mu (\not{p}_1 - \not{k}) \gamma^\nu \not{p}_1 \gamma^\mu (\not{q} + \not{p}_1) \gamma_\nu \right], \quad (122)$$

$$\sum_{\text{spins / pols.}} |\mathcal{M}_s^* \mathcal{M}_t| = \frac{e^2 g^2 C_F N_C}{st} \mu^{4-D} \text{Tr} \left[ \not{p}_2 \gamma^\nu (\not{q} + \not{p}_1) \gamma^\mu \not{p}_1 \gamma^\nu (\not{p}_1 - \not{k}) \gamma_\mu \right]. \quad (123)$$



After some Dirac algebra and plugging in the Mandelstam variables from (113) - (116) this gives

$$\begin{aligned} \sum_{\text{spins / polys.}} |\mathcal{M}_t|^2 &= \frac{e^2 g^2 C_F N_C}{t^2} \mu^{4-D} \left[ 8(D-2)^2 (p_1 \cdot k)(p_2 \cdot k) \right] \\ &= -2 \frac{e^2 g^2 C_F N_C}{t^2} \mu^{4-D} (D-2) \left[ (D-2)st \right], \end{aligned} \quad (124)$$

$$\begin{aligned} \sum_{\text{spins / polys.}} |\mathcal{M}_s|^2 &= \frac{e^2 g^2 C_F N_C}{t^2} \mu^{4-D} \left[ 8(D-2)^2 (q \cdot p_2)(q \cdot p_1) \right] \\ &= -2 \frac{e^2 g^2 C_F N_C}{s^2} \mu^{4-D} (D-2) \left[ (D-2)st \right], \end{aligned} \quad (125)$$

$$\begin{aligned} \sum_{\text{spins / polys.}} |\mathcal{M}_t^* \mathcal{M}_s| &= \frac{e^2 g^2 C_F N_C}{st} \mu^{4-D} 4(D-2) \left[ D(q \cdot k)(p_1 \cdot p_2) - D(q \cdot p_2)(p_1 \cdot k) \right. \\ &\quad \left. - D(q \cdot p_1)(p_2 \cdot k) + 4(q \cdot p_1)(p_1 \cdot p_2) - 8(q \cdot k)(p_1 \cdot p_2) \right. \\ &\quad \left. + 4(q \cdot p_2)(p_1 \cdot k) - 4(p_1 \cdot p_2)(p_1 \cdot k) + 4(q \cdot p_1)(p_2 \cdot k) \right] \\ &= -2 \frac{e^2 g^2 C_F N_C}{st} \mu^{4-D} (D-2) \left[ (D-4)st + 2uq^2 \right], \end{aligned} \quad (126)$$

$$\begin{aligned} \sum_{\text{spins / polys.}} |\mathcal{M}_s^* \mathcal{M}_t| &= \frac{e^2 g^2 C_F N_C}{st} \mu^{4-D} 4(D-2) \left[ D(q \cdot k)(p_1 \cdot p_2) - D(q \cdot p_2)(p_1 \cdot k) \right. \\ &\quad \left. - D(q \cdot p_1)(p_2 \cdot k) + 4(q \cdot p_1)(p_1 \cdot p_2) - 8(q \cdot k)(p_1 \cdot p_2) \right. \\ &\quad \left. + 4(q \cdot p_2)(p_1 \cdot k) - 4(p_1 \cdot p_2)(p_1 \cdot k) + 4(q \cdot p_1)(p_2 \cdot k) \right] \\ &= -2 \frac{e^2 g^2 C_F N_C}{st} \mu^{4-D} (D-2) \left[ (D-4)st + 2uq^2 \right]. \end{aligned} \quad (127)$$

Adding everything back together and averaging over initial states we get in  $D = 4 - 2\epsilon$  dimensions:

$$\frac{1}{4N_F} \sum_{\text{spins / polys.}} |\mathcal{M}|^2 = 2e^2 g^2 C_F \mu^{4-D} (1-\epsilon) \left[ (1-\epsilon) \left( -\frac{s}{t} - \frac{t}{s} \right) - \frac{2uq^2}{st} + 2\epsilon \right], \quad (128)$$

which we will denote with just  $|\mathcal{M}|^2$  from now on for simplicity. In addition to  $|\mathcal{M}|^2$  we need the final particles phase space

$$\int d\Pi_{\text{LIPS}} = \int \frac{d^D p_2}{(2\pi)^{D-1}} \int \frac{d^D k}{(2\pi)^{D-1}} \frac{1}{2E_{p_2}} \frac{1}{2E_k} (2\pi)^D \delta^{(D)}(q + p_1 - p_2 - k). \quad (129)$$

In the c.m.s. of the parton and the virtual photon with the momenta along the  $(D-1)$ th direction we can write the momentum of the final state gluon as

$$k = (|k|, \dots, |k| \cos \theta), \quad (130)$$

where the momenta along the other  $(D-2)$  axes are unspecified. We can integrate over  $p_2$  and rewrite the  $k$  integral into spherical coordinates and solve it from there:

$$\begin{aligned}
\int d\Pi_{\text{LIPS}} &= \frac{1}{4E_{p_2}E_k} (2\pi)^{2-D} \int_0^\infty dE_k E_k^{D-2} \int d\Omega_{D-1} \delta(E_q + E_{p_1} - E_{p_2} - E_k) \\
&= \frac{(2\pi)^{2-D}}{4} \int_0^\infty dE_k E_k^{D-3} \int d\Omega_{D-2} d\theta (\sin\theta)^{D-3} \delta(E_{p_2}(E_q + E_{p_1} - E_{p_2} - E_k)) \\
&= \frac{(2\pi)^{2-D}}{4} \frac{2\pi^{\frac{D-2}{2}}}{\Gamma(\frac{D-2}{2})} \int_0^\infty dE_k E_k^{D-3} \int_{-1}^1 d(\cos\theta) (\sin\theta)^{D-4} \delta(\dots). \tag{131}
\end{aligned}$$

In  $D = 4 - 2\epsilon$  dimensions:

$$= \frac{(4\pi)^{\epsilon-1}}{2} \frac{1}{\Gamma(1-\epsilon)} \int_0^\infty dE_k E_k^{1-2\epsilon} \int_{-1}^1 d(\cos\theta) (1 - \cos^2\theta)^{-\epsilon} \delta(\dots). \tag{132}$$

In the c.m.s. with massless particles we have  $E_k = E_{p_2}$  and can rewrite the delta function

$$= \frac{(4\pi)^{\epsilon-1}}{2} \frac{1}{\Gamma(1-\epsilon)} \int_0^\infty dE_k E_k^{1-2\epsilon} \int_{-1}^1 d(\cos\theta) (1 - \cos^2\theta)^{-\epsilon} \delta(2\sqrt{s}E_k - s), \tag{133}$$

so that we can evaluate the  $E_k$  integral:

$$= \frac{(4\pi)^{\epsilon-1}}{2} \frac{1}{\Gamma(1-\epsilon)} \frac{1}{2\sqrt{s}} \left(\frac{\sqrt{s}}{2}\right)^{1-2\epsilon} \int_{-1}^1 d(\cos\theta) (1 - \cos^2\theta)^{-\epsilon}. \tag{134}$$

Finally we want to substitute  $y = \frac{1}{2}(1 + \cos\theta)$ :

$$\begin{aligned}
&= \frac{(4\pi)^{\epsilon-1}}{8} \frac{1}{\Gamma(1-\epsilon)} \left(\frac{4}{s}\right)^\epsilon \int_0^1 dy 2(4y(1-y))^{-\epsilon} \\
&= \frac{1}{16\pi} \left(\frac{4\pi}{s}\right)^\epsilon \frac{1}{\Gamma(1-\epsilon)} \int_0^1 dy (y(1-y))^{-\epsilon}. \tag{135}
\end{aligned}$$

With the c.m.s. Mandelstam variables

$$s = \frac{Q^2}{z}(1-z), \quad t = -\frac{Q^2}{z}(1-y), \quad u = -\frac{Q^2}{z}y, \tag{136}$$

we can use (128) and (135) to get  $\hat{W}_0^{\text{R}}$ :

$$\begin{aligned}
e^2 \hat{W}_0^{\text{R}} &= 4\pi e_i^2 \int |\mathcal{M}|^2 d\Pi_{\text{LIPS}} \\
&= 4\pi e_i^2 \frac{e^2 \alpha_s}{2\pi} C_F \left(\frac{4\pi\mu^2}{Q^2}\right)^\epsilon \frac{1}{\Gamma(1-\epsilon)} \left\{ 3z + z^\epsilon(1-z)^{-\epsilon} \int_0^1 dy (y(1-y))^{-\epsilon} \right. \\
&\quad \left. \left[ (1-\epsilon) \left( \frac{1-z}{1-y} + \frac{1-y}{1-z} \right) + \frac{2zy}{(1-y)(1-z)} \right] \right\}. \tag{137}
\end{aligned}$$

This integral can be evaluated for  $\epsilon > 0$ :

$$\begin{aligned}\mathcal{I}(z) &= \int_0^1 dy (y(1-y))^{-\epsilon} \left[ (1-\epsilon) \left( \frac{1-z}{1-y} + \frac{1-y}{1-z} \right) + \frac{2zy}{(1-y)(1-z)} \right] \\ &= 4^{-1-\epsilon} \sqrt{\pi} \frac{\Gamma(\epsilon)}{(z-1)\Gamma(\frac{3}{2}+\epsilon)} \left( -8z\epsilon^2 + 2z^2(\epsilon-1)(1+2\epsilon) + (\epsilon-1)(2+5\epsilon) \right).\end{aligned}\quad (138)$$

When plugging this back into (137) we pull a factor of  $\frac{\Gamma(1-\epsilon)^2}{\Gamma(1-2\epsilon)}$  out of the parentheses and expand the inside as a power series in  $\epsilon$ :

$$\begin{aligned}\hat{W}_0^R &= 4\pi e_i^2 \frac{\alpha_s}{2\pi} C_F \left( \frac{4\pi\mu^2}{Q^2} \right)^\epsilon \frac{\Gamma(1-\epsilon)}{\Gamma(1-2\epsilon)} \\ &\quad \times \left\{ 3z + \frac{z^\epsilon}{(1-z)^{1+\epsilon}} \left( -\frac{1}{\epsilon}(1+z^2) + (3-z)(1-z) - \frac{3}{2} - \frac{7}{2}\epsilon \right) \right\}.\end{aligned}\quad (139)$$

The  $\frac{1}{\epsilon^2}$  double pole in  $\hat{W}_0^R$  is hidden inside the  $(1-z)^{-1-\epsilon}$ . We can expand this term around  $\epsilon = 0$  using the plus-distribution identity (A.37) that is derived in the section A.1.5:

$$\begin{aligned}\hat{W}_0^R &= 4\pi e_i^2 \frac{\alpha_s}{2\pi} C_F \left( \frac{4\pi\mu^2}{Q^2} \right)^\epsilon \frac{\Gamma(1-\epsilon)}{\Gamma(1-2\epsilon)} \\ &\quad \times \left\{ 3z + z^\epsilon \left[ (1+z^2) \left( \frac{1}{\epsilon^2} \delta(1-z) - \frac{1}{\epsilon} \left[ \frac{1}{1-z} \right]_+ + \left[ \frac{\ln(1-z)}{1-z} \right]_+ \right) \right. \right. \\ &\quad \left. \left. + 3(1-z)^{-\epsilon} - z(1-z)^{-\epsilon} + \frac{3}{2} \frac{1}{\epsilon} \delta(1-z) - \frac{3}{2} \left[ \frac{1}{1-z} \right]_+ + \frac{7}{2} \delta(1-z) \right] \right\}.\end{aligned}\quad (140)$$

Expanding everything into powers of  $\epsilon$  (up to  $\mathcal{O}(\epsilon^0)$ ):

$$\begin{aligned}\hat{W}_0^R &= 4\pi e_i^2 \frac{\alpha_s}{2\pi} C_F \left( \frac{4\pi\mu^2}{Q^2} \right)^\epsilon \frac{\Gamma(1-\epsilon)}{\Gamma(1-2\epsilon)} \\ &\quad \times \left\{ 3z + \delta(1-z) \left[ \frac{1}{\epsilon^2}(1+z^2) \left( 1 + \epsilon \ln z + \frac{1}{2} \epsilon^2 \ln^2 z \right) + \frac{3}{2} \frac{1}{\epsilon} (1 + \epsilon \ln z) + \frac{7}{2} \right] \right. \\ &\quad \left. - \frac{1+z^2}{\epsilon} (1 + \epsilon \ln z) \left[ \frac{1}{1-z} \right]_+ + (1+z^2) \left[ \frac{\ln(1-z)}{1-z} \right]_+ - \frac{3}{2} \left[ \frac{1}{1-z} \right]_+ + 3 - z \right\}.\end{aligned}\quad (141)$$

After taking advantage of the delta distribution  $\delta(1-z)$  to simplify some terms we get to the final result:

$$\begin{aligned}\hat{W}_0^R &= 4\pi e_i^2 \frac{\alpha_s}{2\pi} C_F \left( \frac{4\pi\mu^2}{Q^2} \right)^\epsilon \frac{\Gamma(1-\epsilon)}{\Gamma(1-2\epsilon)} \left\{ 3 + 2z - \frac{1+z^2}{1-z} \ln z + \left( \frac{2}{\epsilon^2} + \frac{3}{2\epsilon} + \frac{7}{2} \right) \delta(1-z) \right. \\ &\quad \left. - \left( \frac{1+z^2}{\epsilon} + \frac{3}{2} \right) \left[ \frac{1}{1-z} \right]_+ + (1+z^2) \left[ \frac{\ln(1-z)}{1-z} \right]_+ \right\}.\end{aligned}\quad (142)$$

### 2.7.2 DGLAP equations

Having calculated all necessary diagrams, we can now add the leading order value and the two corrections (109), (142) together to get the total  $\hat{W}_0$  at next-to-leading order:

$$\begin{aligned}\hat{W}_0 &= \hat{W}_0^{\text{LO}} + \hat{W}_0^V + \hat{W}_0^R \\ &= 4\pi e_i^2 \left\{ \left[ \delta(1-z) - \frac{1}{\epsilon} \frac{\alpha_s}{2\pi} P_{qq}(z) \left( \frac{4\pi\mu^2}{Q^2} \right)^\epsilon \frac{\Gamma(1-\epsilon)}{\Gamma(1-2\epsilon)} \right] + \frac{\alpha_s}{2\pi} C_F \left[ -\frac{3}{2} \left[ \frac{1}{1-z} \right]_+ \right. \right. \\ &\quad \left. \left. + (1+z^2) \left[ \frac{\ln(1-z)}{1-z} \right]_+ - \frac{1+z^2}{1-z} \ln z + 3 + 2z - \left( \frac{9}{2} + \frac{1}{3}\pi^2 \right) \delta(1-z) \right] \right\},\end{aligned}\quad (143)$$

with

$$P_{qq}(z) = C_F \left[ (1+z^2) \left[ \frac{1}{1-z} \right]_+ + \frac{3}{2} \delta(1-z) \right]. \quad (144)$$

This  $P_{qq}(z)$  is called a splitting function. The remaining single  $\frac{1}{\epsilon}$  pole has a residue proportional to  $P_{qq}(z)$ . Having a pole in the parton-level cross section is not a problem because it is not a physical observable. If we plug (143) into (97) we get

$$W_0(x, Q^2) = 4\pi \sum_i e_i^2 \int_x^1 \frac{d\xi}{\xi} f_i(\xi) \left[ \delta\left(1 - \frac{x}{\xi}\right) - \frac{\alpha_s}{2\pi} P_{qq}\left(\frac{x}{\xi}\right) \left( \frac{1}{\epsilon} + \ln \frac{\tilde{\mu}^2}{Q^2} \right) + \text{finite} \right], \quad (145)$$

for the physical cross section. Since the splitting function satisfies

$$\int_0^1 P_{qq}(z) dz = 0, \quad (146)$$

the pole vanishes and the total DIS cross section at any given  $Q^2$  is finite, when integrated over  $x$ :

$$\begin{aligned}\int_0^1 dx \int_x^1 \frac{d\xi}{\xi} f_i(\xi) P_{qq}\left(\frac{x}{\xi}\right) &= \int_0^1 dx \int_0^1 d\xi \int_0^1 dz f_i(\xi) P_{qq}\left(\frac{x}{\xi}\right) \delta(x - \xi z) \\ &= \int_0^1 d\xi \int_0^1 dz f_i(\xi) P_{qq}(z) \\ &= \int_0^1 d\xi f_i(\xi) \int_0^1 dz P_{qq}(z) = 0.\end{aligned}\quad (147)$$

For any fixed  $x$  the cross section is divergent. Thus we have to take differences at different scales  $Q^2$  and  $Q_0^2$  to get a finite result:

$$W_0(x, Q^2) - W_0(x, Q_0^2) = 4\pi \sum_i e_i^2 \int_x^1 \frac{d\xi}{\xi} f_i(\xi) \left[ \frac{\alpha_s}{2\pi} P_{qq}\left(\frac{x}{\xi}\right) \ln \frac{Q^2}{Q_0^2} \right]. \quad (148)$$

The  $\frac{1}{\epsilon}$  pole in the parton-level cross section (143) therefore leads to a physical prediction of logarithmic  $Q^2$  dependence in the hadronic cross section. The reason why we need to take differences here, even though  $W_0(x, Q^2)$  should be observable by itself (and therefore finite), is that we did not do the calculation in full QCD where the divergence would be cut off by a

physical scale such as the quark mass. The same can be found in QED where an equivalent divergence in Compton scattering gets cut off by the electron mass. However, this calculation is still valuable as it gives us a testable prediction from perturbation theory, which fits well with the violation of Bjorken scaling seen in figure 2.9 that we set out to explain in this chapter.

Through renormalization we can replace the calculation of differences by the calculation of observables in terms of renormalized quantities. This is done by defining some reference scale  $Q_0^2$  at which  $W_0$  is given exactly by (102):

$$W_0(x, Q^2) \equiv 4\pi \sum_i e_i^2 f_i(x, \mu^2 = Q^2). \quad (149)$$

The renormalized PDFs have to fulfill

$$f_i(x, \mu_1^2) = f_i(x, \mu^2) + \frac{\alpha_s}{2\pi} \int_x^1 \frac{d\xi}{\xi} f_i(\xi, \mu_1^2) P_{qq}\left(\frac{x}{\xi}\right) \ln \frac{\mu_1^2}{\mu^2}, \quad (150)$$

which leads to the DGLAP evolution equation

$$\mu^2 \frac{d}{d\mu^2} f_i(x, \mu^2) = \frac{\alpha_s}{2\pi} \int_x^1 \frac{d\xi}{\xi} f_i(\xi, \mu^2) P_{qq}\left(\frac{x}{\xi}\right). \quad (151)$$

One important check of our results is that  $f_i(x)$  should fulfill sum rules like (79) at any scale  $\mu^2$ . If we integrate (150) and use (147) like we did before, we see that  $\int f_i(x, \mu^2) dx$  is indeed scale invariant.

Until now we have not considered any processes unrelated to  $e^- p^+ \rightarrow e^- X$ , where quarks are the only partons in the initial state. At next-to-leading order there are also processes like  $\gamma^* g \rightarrow q \bar{q}$  with gluons in the initial state. This leads to more splitting functions and mixing of the various PDFs under the full evolution equation:

$$\mu^2 \frac{d}{d\mu^2} \begin{pmatrix} f_i(x, \mu^2) \\ f_g(x, \mu^2) \end{pmatrix} = \sum_j \frac{\alpha_s}{2\pi} \int_x^1 \frac{d\xi}{\xi} \begin{pmatrix} P_{q_i q_j}\left(\frac{x}{\xi}\right) & P_{q_i g}\left(\frac{x}{\xi}\right) \\ P_{g q_j}\left(\frac{x}{\xi}\right) & P_{g g}\left(\frac{x}{\xi}\right) \end{pmatrix} \begin{pmatrix} f_j(\xi, \mu^2) \\ f_g(\xi, \mu^2) \end{pmatrix}. \quad (152)$$

The other splitting functions can be derived from the cross sections of other processes in a similar manner to the way we derived  $P_{qq}(z)$ . At leading order they are given by [28]

$$P_{qq}(z) = C_F \left[ \frac{1+z^2}{[1-z]_+} + \frac{3}{2} \delta(1-z) \right], \quad (153)$$

$$P_{qg}(z) = T_F \left[ z^2 + (1-z)^2 \right], \quad (154)$$

$$P_{gq}(z) = C_F \left[ \frac{1+(1-z)^2}{z} \right], \quad (155)$$

$$P_{gg}(z) = 2C_A \left[ \frac{z}{[1-z]_+} + \frac{1-z}{z} + z(1-z) \right] + \delta(1-z) \left[ \frac{11}{6} C_A - \frac{2}{3} T_F n_f \right]. \quad (156)$$

Higher order contributions to the splitting functions have been calculated and can be found for example in [29], [30] (NLO) and [31], [32] (NNLO).

### 2.7.3 Scale dependence

Looking at equation (152) does not grant much insight about the effect that the DGLAP evolution will have on actual PDFs. Therefore we will take the nCTEQ15 PDFs [33] and evolve them to various scales numerically using HOPPET [34]. Figure 2.11 shows the valence quark distributions  $u_v = u - \bar{u}$  and  $d_v = d - \bar{d}$  while figure 2.10 shows the distribution of each flavour, gluons and the antiquarks for  $u$  and  $d$  (of course for all other flavours the distributions for quarks and antiquarks are the same).

We can see the gluon distribution in the bottom right graph of figure 2.10 which starts at the initial scale as a mostly flat curve with a small peak around  $x \approx 0.02$ . The region around  $x = 0.05$  hardly gets affected by the scale, while the region above gets slightly suppressed at higher scales. The small  $x$  region rises rapidly as the scale increases and due to the splitting of the gluons all sea quark distributions can be seen to rise in the same region in the same shape. The first two graphs in figure 2.10 show the seaquarks  $\bar{u}$  and  $\bar{d}$ . As expected they look very similar overall but they differ somewhat in the region around  $x = 0.1$  at low energy  $Q^2$  due to the differences in mass. In the small  $x$  region these two graphs look basically indistinguishable from each other and also the  $u$  and  $d$  distributions because the valence quarks hardly make a contribution in this region. Conversely in the large  $x$  region the  $u$  and  $d$  distributions are almost entirely determined by the valence quarks. Lastly, the  $s$ ,  $c$  and  $b$  graphs all look rather similar at higher scales where the difference in mass becomes less and less important. At the initial scale,  $s$  is the only flavour among the three that has a non-zero distribution in the chosen parameterization. Qualitatively its shape looks very similar to the  $\bar{u}$  and  $\bar{d}$  distributions albeit somewhat lower due to the heavier mass of the strange quark. Strange and charm flavours get almost identical contributions from the evolution and would look very alike if one were to subtract the initial  $s$  distribution at every scale. The bottom quark on the other hand hardly gets a contribution from the evolution at the second lowest scale  $Q^2 = 6.25 \text{ GeV}^2$  as this still lies in the region of the bottom quarks mass and only slowly catches up to the other two flavours at higher energies.

The two valence quark distributions are hardly affected by the evolution in the small  $x$  region as the contributions from the sea  $u$  and  $\bar{u}$  cancel out. In the medium and large  $x$  region we see the valence quark contributions slightly diminished on higher scales, meaning they carry a smaller fraction of the hadron's overall momentum. Accordingly the strong rise in the gluon's small  $x$  region at higher scales means that they take over this portion of the momentum.

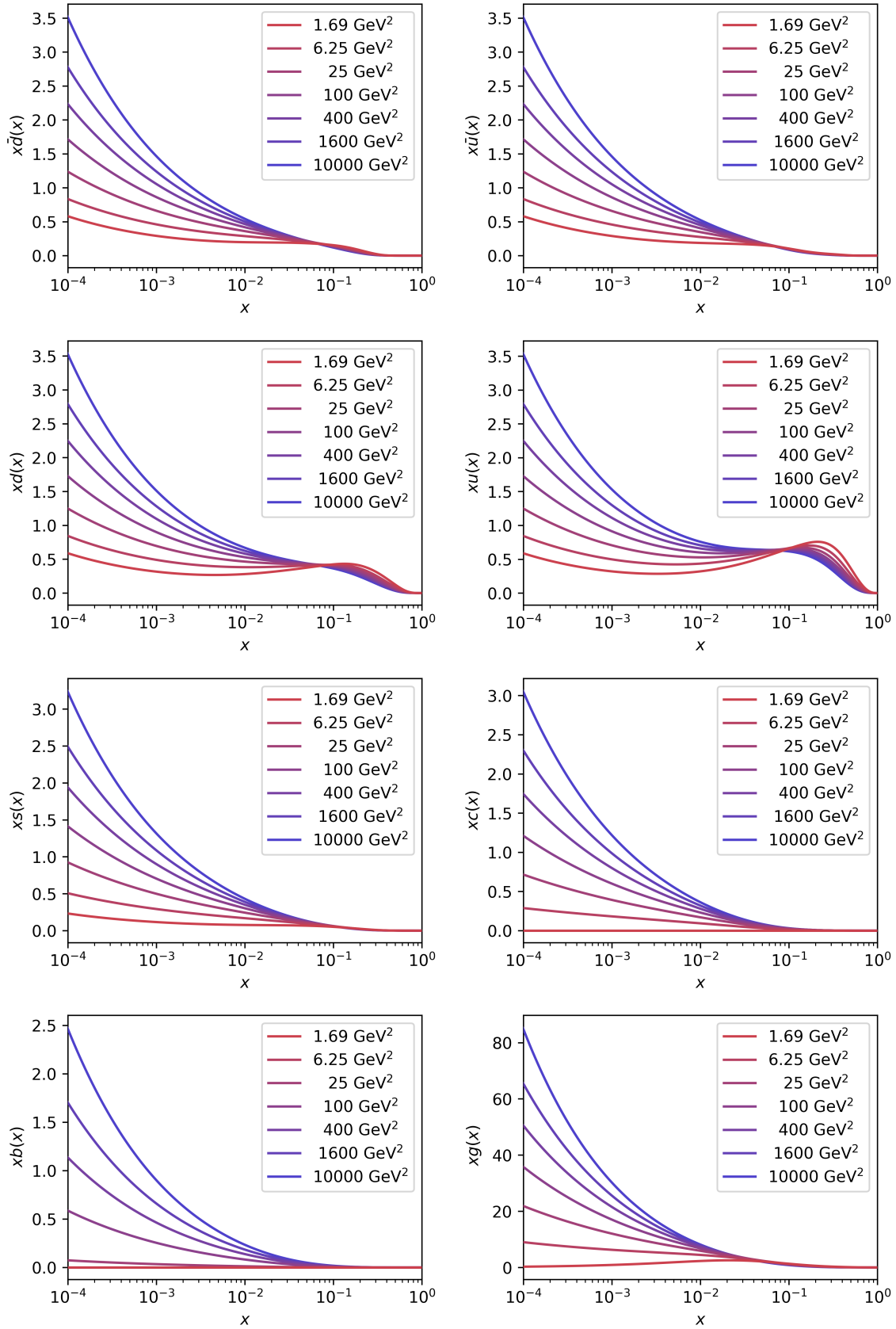


Figure 2.10: Parton distribution functions  $xf(x)$  at various scales  $Q^2$  from the input scale  $Q_0^2 = 1.69 \text{ GeV}^2$  to  $Q_0^2 = 10^4 \text{ GeV}^2$ .

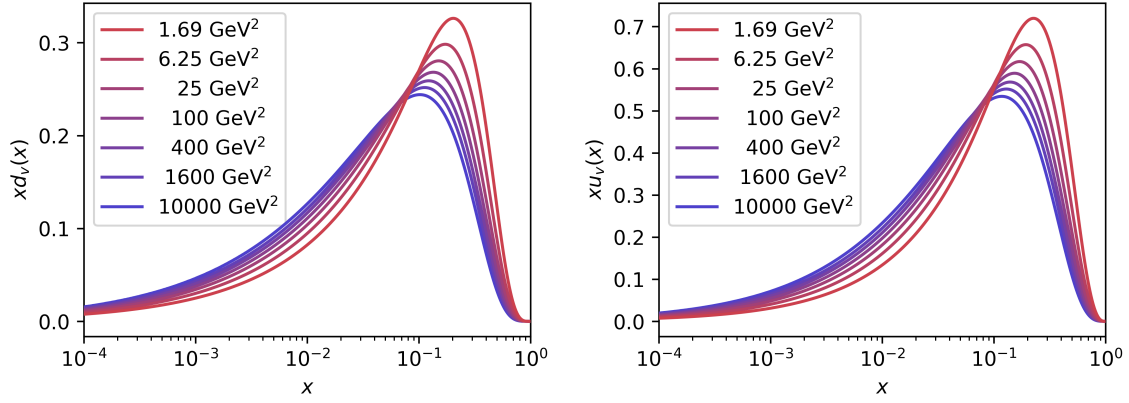


Figure 2.11: Parton distribution functions  $xf(x)$  for the valence quarks at various scales  $Q^2$  from the input scale  $Q_0^2 = 1.69 \text{ GeV}^2$  to  $Q_0^2 = 10^4 \text{ GeV}^2$ .

### 3 Modelling of data

Another main goal of this thesis is to find an approximation of the non-perturbative parton distribution functions from experimental data. To do this, we will choose parameterizations for the various types of PDFs (e.g.  $u, d, g, \bar{u} + \bar{d}$ ) and then search for the best possible set of parameters to match experimental data.

Choosing the right parameterization itself is an important and non-trivial first step. The chosen function needs to strike a good balance between two opposing qualities:

1. It needs to be flexible enough to be able to match the real function as accurately as possible.
2. It needs to have as few parameters as possible, so that they can be determined with high certainty.

The problem of quantifying the uncertainties however shall be taken care of later. For now we want to understand how to find the right parameters given a set of  $N$  data points  $(x_i, y_i), i \in [0, \dots, N-1]$  and a function  $f(x; \mathbf{a})$  depending on a set of parameters  $a_j, j \in [0, \dots, M-1]$  and a variable  $x$  (the generalization from one variable  $x$  to multiple input variables is trivial).

#### 3.1 Least squares fitting

Once we have a functional model of our data set we need a merit-function that measures how well the function with a particular set of parameters matches the data. We can get such a merit function from maximum likelihood estimation, which asks the question of how likely the measured set of data would occur given a particular set of parameters. Of course the probability would always be zero, if  $y$  takes on continuous values, unless we allow some fixed deviation  $\pm \Delta y$  at each data point. If the data set is very improbable given the set of parameters we expect these parameters to be unlikely too. Therefore we use the probability of the data given the



parameters as a measure of the likelihood of the parameters according to the data. The best-fit parameters are therefore the set of parameters that maximizes the above probability. Assuming normal distributed, independent measurement errors for each point, the probability of the data set is just the product of the probabilities of each point:

$$P \propto \prod_{i=0}^{N-1} \exp \left[ -\frac{1}{2} \left( \frac{y_i - y(x_i; \mathbf{a})}{\sigma_i} \right)^2 \right] \Delta y. \quad (157)$$

Maximizing this product is the same as minimizing its negative logarithm:

$$\left( \sum_{i=0}^{N-1} \frac{[y_i - y(x_i; \mathbf{a})]^2}{2\sigma_i^2} \right)^2 - N \log \Delta y. \quad (158)$$

Dropping the constant term gives us

$$\chi^2(\mathbf{a}) \equiv \left( \sum_{i=0}^{N-1} \frac{[y_i - y(x_i; \mathbf{a})]^2}{2\sigma_i^2} \right)^2. \quad (159)$$

Minimizing this function (159) is called weighted least squares fitting or chi-square fitting. The term “weighted” comes from the fact that we assume different standard deviations  $\sigma_i$  for each point.

At this point we need to make some remarks regarding the validity of our assumptions. In reality it is often the case that the data is not actually normally distributed around the true model. A simple example are experiments where events are counted. The errors are usually distributed according to Poisson statistics, which converges to a Gaussian, if the number of events is large enough. The problem with this is, that the convergence is not uniform. When going further away from the mean of the distribution one needs an increasing amount of events for the distribution to converge to a Gaussian. Since the normal distribution predicts less events on the “tail” of the distribution than actually occur according to Poisson, the least-squares fit gets skewed towards these events. Also worth mentioning, although rare, are outliers. Outliers are data points very far away from the true model usually due to some external influence. These outliers can have very strong negative effects on least squares fits, because their probability in the Gaussian model is so small that the maximum likelihood estimation might heavily distort the curve to account for them. Aside from the statistical errors that vanish when taking enough data, there are also systematic errors, like slight miscalibration of measurement devices that cannot be mitigated by any averaging or numerical processing.

In the case of ideal data, where there are no systematic errors and all statistical errors are actually normally distributed, the  $\chi^2$  distribution has a mean equal to the number of degrees of freedom  $\nu = N_{\text{data}} - N_{\text{parameters}}$  with a standard deviation  $\sqrt{2\nu}$ . This means we can estimate the  $\chi^2$  value of an acceptable fit to be roughly  $\nu$ .

### Correlated uncertainties

The simple definition of  $\chi^2$  given by (159) is not always sufficient when working with real data. By modifying the definition of  $\chi^2$  as derived in [35] we can take correlations in the experimental

uncertainties into account. To do this we first split  $\chi^2$  into the values  $\chi_n^2$  for each experiment:

$$\chi^2(\mathbf{a}) = \sum_n w_n \chi_n^2(\mathbf{a}), \quad (160)$$

where  $w_n$  are weighting factors that we will leave at a value of 1. The correlated errors are denoted by  $\beta_{ik}$  where the indices  $i = 1, 2, \dots, N_{\text{data}}$  are for the data points and  $k = 1, 2, \dots, K$  for the sources of correlated errors. We then define a vector  $B$  and a matrix  $A$

$$B_k \equiv \sum_{i=1}^{N_{\text{data}}} \frac{\beta_{ki}(D_i - T_i)}{\hat{\sigma}_i^2} \quad \text{and} \quad A_{kk'} \equiv \delta_{kk'} + \sum_{i=1}^{N_{\text{data}}} \frac{\beta_{ki}\beta_{k'i}}{\hat{\sigma}_i^2}, \quad (161)$$

where  $\hat{\sigma}_i^2$  is given by the sum of the squares of the statistical errors  $\sigma_i$  and uncorrelated systematic errors  $u_i$ :

$$\hat{\sigma}_i^2 \equiv \sigma_i^2 + u_i^2. \quad (162)$$

The contribution from each experiment is then given by

$$\chi_n^2(\mathbf{a}) = \sum_i^{N_{\text{data}}} \frac{[D_i - T_i(\mathbf{a})]^2}{\hat{\sigma}_i^2} - \sum_{k,k'}^K B_k A_{k,k'}^{-1} B_{k'}. \quad (163)$$

This correction to the  $\chi^2$  function allows us to account for all experimental uncertainties, but theoretical uncertainties from missing higher order corrections or choice of parametrization of the PDFs are not included.

### 3.2 Nonlinear least squares fitting

The most general models depend nonlinearly on a set of  $M$  unknown parameters  $a_i, i \in [0, \dots, M-1]$ . If the function depends only linearly on the parameters the best fit can be found in a single step by using simple linear regression, while functions with nonlinear dependence on the parameters require an iterative procedure. There are two ways to go about this: Sufficiently close to the minimum of  $\chi^2$  it can be approximated as

$$\chi^2(\mathbf{a}) \approx \gamma - \mathbf{d} \cdot \mathbf{a} + \frac{1}{2} \mathbf{a} \cdot \mathbf{H} \cdot \mathbf{a}, \quad (164)$$

where  $\mathbf{d}$  is an  $M$ -vector and  $\mathbf{H}$  an  $M \times M$  matrix. In this case the best fit parameters  $a_{\text{min}}$  can be obtained from the current set  $a_{\text{cur}}$  by

$$\mathbf{a}_{\text{min}} = \mathbf{a}_{\text{cur}} + \mathbf{H}^{-1} \cdot [-\nabla \chi^2(\mathbf{a}_{\text{cur}})]. \quad (165)$$

This is referred to as the inverse Hessian method.

Far from the minimum the approximation (164) might not work well and it is better to use the

steepest descend method to get closer to the minimum:

$$\mathbf{a}_{\text{next}} = \mathbf{a}_{\text{cur}} - c \cdot \nabla \chi^2(\mathbf{a}_{\text{cur}}), \quad (166)$$

where  $c$  is a small constant preventing  $\mathbf{a}_{\text{next}}$  from overshooting  $\mathbf{a}_{\text{min}}$ . Both ways depend on the gradient of  $\chi^2$  and the inverse Hessian method also depends on the matrix of second derivatives. The gradient and the matrix of second derivatives  $\mathbf{H}$  can be calculated exactly if the model function is known in closed form. This will not be the case for PDFs however, since the parametrization is only known at the initial scale  $Q_0$  and its form after DGLAP evolution is unknown. We will use forward finite differences to approximate the derivatives because they only depend on one additional function evaluation per parameter and each evaluation with new parameters requires another computationally expensive DGLAP evolution to be performed. The relatively low accuracy of the forward finite difference does not affect the final value of the best fit  $\chi^2$  but might increase the number of iterations needed to reach it.

The gradient of the merit function (159) with respect to the parameters is given by

$$\frac{\partial \chi^2}{\partial a_k} = -2 \sum_{i=0}^{N-1} \frac{y_i - y(x_i; \mathbf{a})}{\sigma_i^2} \frac{\partial y(x_i, \mathbf{a})}{\partial a_k}, \quad \text{with } k \in [0, M-1]. \quad (167)$$

The second derivative is then

$$\frac{\partial^2 \chi^2}{\partial a_k \partial a_l} = 2 \sum_{i=0}^{N-1} \frac{1}{\sigma^2} \left( \frac{\partial y(x_i, \mathbf{a})}{\partial a_k} \frac{\partial y(x_i, \mathbf{a})}{\partial a_l} - [y_i - y(x_i; \mathbf{a})] \frac{\partial^2 y(x_i, \mathbf{a})}{\partial a_k \partial a_l} \right). \quad (168)$$

To remove the factors before the sum we define

$$\beta_k \equiv -\frac{1}{2} \frac{\partial \chi^2}{\partial a_k} \quad \text{and} \quad \alpha_{kl} \equiv \frac{1}{2} \frac{\partial^2 \chi^2}{\partial a_k \partial a_l}. \quad (169)$$

When talking about the Hessian from here on we will always mean the matrix  $\alpha$ , as it is convention in high energy physics. With  $\mathbf{H} = 2\alpha$  we can rewrite (165) as

$$\sum_{l=0}^{M-1} \alpha_{kl} \delta a_l = \beta_k, \quad (170)$$

where  $\delta a_l$  is the shift of parameter  $l$  from the current approximation to the new one. The steepest descent method can be expressed as

$$\delta a_l = c \cdot \beta_l. \quad (171)$$

The Hessian depends both on the first and second derivatives of the model function with respect to the parameters. The second derivatives are often ignored because close to the minimum, they should be smaller than the term involving the first derivatives. Another argument for ignoring them is that the term  $(y_i - y(x_i; \mathbf{a}))$  in front of the second derivative should be equal to the random noise if the model is accurate enough and hence it should cancel out when summed

over  $i$ . Therefore we will use

$$\alpha_{kl} = \sum_{i=0}^{N-1} \frac{1}{\sigma_i^2} \left( \frac{\partial y(x_i; \mathbf{a})}{\partial a_k} \frac{\partial y(x_i; \mathbf{a})}{\partial a_l} \right) \quad (172)$$

as the definition of the Hessian in our program.

### 3.2.1 Levenberg-Marquardt method

The central idea of the Levenberg-Marquardt method is a smooth variation between the inverse Hessian and the steepest descent method. There are two more considerations necessary to understand this method. The first question is how the constant  $c$  in (171) should be chosen. The Hessian can give us insight into this magnitude because its elements are the curvature of  $\chi^2$  in parameter space and therefore give us a measure of how far the slope  $\beta$  extends. We can find which elements of the Hessian we need to use from dimensional analysis of the equation:  $\delta a_k$  is the shift in parameter  $a_k$ , so it has to have the same unit  $[a_k]$ . The definition of the vector  $\beta$  then tells us that  $\beta_k$  has the dimension  $[\frac{1}{a_k}]$  since  $\chi^2$  is nondimensional. Therefore the dimension of the constant has to be  $[c] = [a_k^2]$ , which is fulfilled by the reciprocals of the Hessian's diagonal elements  $\frac{1}{\alpha_{kk}}$ . It is necessary that all  $\alpha_{kk}$  are positive, otherwise the gradient descent would go in the wrong direction. This is generally not an issue, however, since (172) guarantees that they are positive. An additional smallness factor  $\frac{1}{\lambda}$  with  $\lambda \gg 1$  is introduced to make sure the steps along the gradient are not too large. Equation (171) then becomes

$$\delta a_l = \frac{1}{\lambda \alpha_{ll}} \beta_l \quad \text{or} \quad \lambda \alpha_{ll} \delta a_l = \beta_l, \quad (173)$$

which is essentially the same form as the definition of the inverse Hessian method (170). This allows us to combine the two into

$$\sum_{l=0}^{M-1} \hat{\alpha}_{kl} \delta a_l = \beta_k, \quad (174)$$

with the new matrix  $\hat{\alpha}$ :

$$\hat{\alpha}_{ll} \equiv \alpha_{ll}(1 + \lambda), \quad (175)$$

$$\hat{\alpha}_{lk} \equiv \alpha_{lk}, \quad (l \neq k). \quad (176)$$

If  $\lambda$  is large the matrix  $\hat{\alpha}$  becomes diagonally dominant, such that the steepest descent is weighted more than the inverse Hessian. For small  $\lambda$  we get  $\hat{\alpha} \approx \alpha$  and we restore equation (170) for the inverse Hessian method. The main algorithm is shown in figure 3.1.

The last missing piece of the algorithm is a stopping condition. As stated before we expect  $\chi^2$  to have its minimum at roughly  $\chi^2 = \nu \pm \sqrt{2\nu}$ . Therefore we stop the algorithm once  $\chi^2$  has reached this region and  $\Delta\chi^2 \ll \sqrt{2\nu}$  in further iterations.

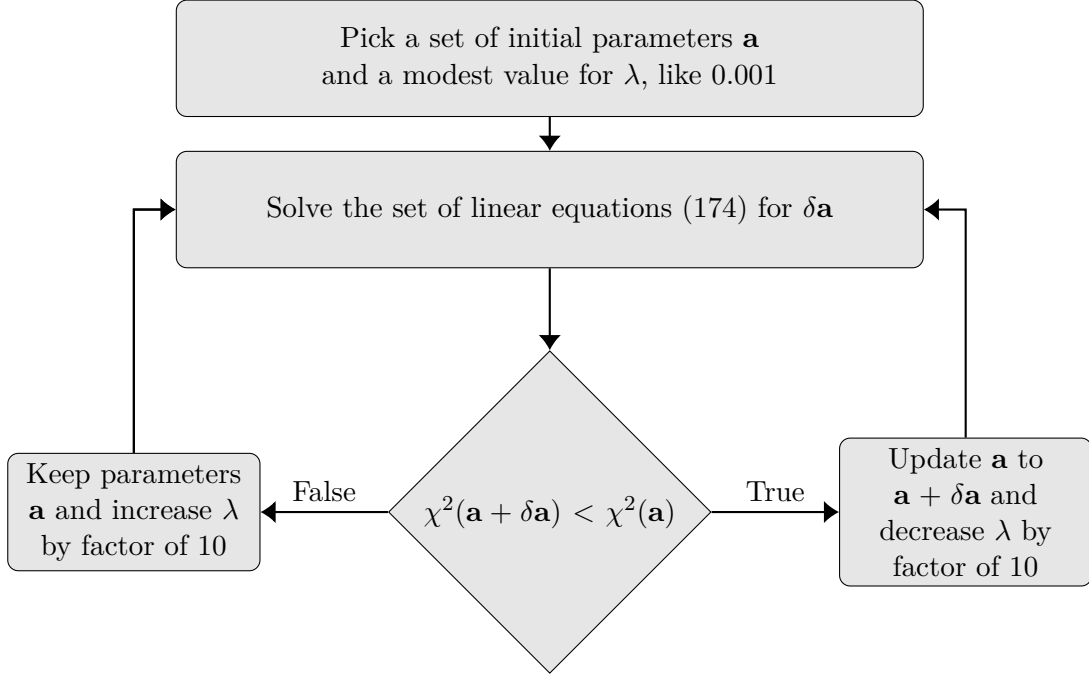


Figure 3.1: Diagrammatic representation of the Levenberg-Marquardt algorithm.

### 3.2.2 Implementation in C++

The implementation of the algorithm is done in C++ using the `armadillo` library [36] to adapt the code from *Numerical Recipes in C++* [37] into a more efficient and readable version for PDFs. Unless otherwise specified, any code from this point on will use the namespaces `std` and `arma`. The main program loop repeatedly calls the function `mrqmin` to perform one iteration of the Levenberg-Marquardt algorithm until  $\chi^2$  is sufficiently small and decreases only by a negligible amount in further iterations. The algorithm should also stop if `lambda` grows too large (e.g.  $10^{20}$ ), signaling bad convergence. On the first call  $\lambda$  is set to a negative value to start the initial setup. For the last call  $\lambda$  is set to 0 and  $\alpha$  is set to the Hessian matrix without any modifications, after being used as a workspace variable during the rest of the runtime.

Each iteration of `mrqmin` calls the function `mrqcof` to calculate the matrix  $\alpha$ , vector  $\beta$  and  $\chi^2$ . There are some technical differences between this process in the code and in the general Levenberg-Marquardt method, originating from the integration into the `nCTEQ` codebase. The `nCTEQ` code makes use of `HOPPET` [34] for DGLAP evolution, `LHAPDF` [38] for interpolation and external PDFs, `APPLgrid` [39] for grid computations and `ROOT / MINUIT2` [40] for a variety of minimization related functions. Instead of a function and a set of data we use a `PDFBase` object and a `Chi2` object. The specifics of these objects are not relevant here. The `Chi2` holds the information about the experimental data and kinematic cuts. The `PDFBase` object, which is known by the `Chi2` instance holds all information about the parametrization, evolution and interpolation.

Instead of evaluating a function  $f(\mathbf{x}; \mathbf{a})$  and its analytical derivatives  $\frac{\partial f}{\partial a_i}(\mathbf{x}; \mathbf{a})$  to calculate  $\beta$  exactly and  $\alpha$  via equation (172), we need to calculate  $\alpha$  and  $\beta$  directly via finite differences of

$\chi^2$ , because  $f(x, Q)$  is not a known analytical function. To do this we first implement a method `eval_chi` that just returns the value of  $\chi^2$  for a given set of parameters. The variables `param_names`, `params`, `chi2obj` and `PDF_instance` will appear in almost all methods for the same purpose and will therefore be explained one time only.

```
double eval_chi(Chi2 &chi2obj, PDFBase& PDF_instance,
    const std::vector<std::string> param_names, vec params){
/* Evaluates chisq at the given parameters.
Variables:
chi2obj: Instance of the Chi2 object
PDF_instance: Instance of the PDF object
param_names: Names of the parameters in the parameterization
params: Current parameter values*/
PDFBase* PDF = &PDF_instance;
PDF->setParameters(param_names, conv_to<std::vector<double>>::from(params));
return chi2obj.chi2();
}
```

Introducing the notation  $\mathbf{h}_i = h_i \mathbf{e}_i$ , where  $\mathbf{e}_i$  is the unit vector in the  $i$ -th direction and  $h_i$  the step size for the  $i$ -th parameter, we can express the finite differences for  $\alpha$  and  $\beta$  as

$$\beta_i(\mathbf{a}) \approx -\frac{\chi^2(\mathbf{a} + \mathbf{h}_i) - \chi^2(\mathbf{a} - \mathbf{h}_i)}{4h_i}, \quad (177)$$

$$\alpha_{ii}(\mathbf{a}) \approx \frac{\chi^2(\mathbf{a} + \mathbf{h}_i) - 2\chi^2(\mathbf{a}) + \chi^2(\mathbf{a} - \mathbf{h}_i)}{2h_i^2}, \quad (178)$$

$$\alpha_{ij}(\mathbf{a}) = \alpha_{ji}(\mathbf{a}) \approx \frac{1}{4h_i h_j} \left[ \chi^2(\mathbf{a} + \mathbf{h}_i + \mathbf{h}_j) + \chi^2(\mathbf{a} - \mathbf{h}_i - \mathbf{h}_j) + 2\chi^2(\mathbf{a}) - \chi^2(\mathbf{a} + \mathbf{h}_i) - \chi^2(\mathbf{a} + \mathbf{h}_j) - \chi^2(\mathbf{a} - \mathbf{h}_i) - \chi^2(\mathbf{a} - \mathbf{h}_j) \right]. \quad (179)$$

We use central differences as a compromise between runtime and accuracy. After  $\chi^2(\mathbf{a})$  and  $\beta(\mathbf{a})$  have been calculated we can reuse many of the terms so that we get the diagonal elements of  $\alpha$  at basically no cost and we only need to evaluate  $\chi^2(\mathbf{a} \pm \mathbf{h}_i \pm \mathbf{h}_j)$  for each  $\alpha_{ij}(\mathbf{a})$ . We have to make this compromise because each evaluation of  $\chi^2$  takes longer than the entire rest of the fitting process by multiple orders of magnitude<sup>5</sup>. This is due to the fact that only  $f(x, Q_0)$  is an analytical function and each evaluation of  $\chi^2$  requires  $f(x, Q \neq Q_0)$  to be calculated via DGLAP evolution for each data point. Note that (178) no longer guarantees the diagonal elements to be positive and to account for that we need to change the definition of  $\hat{\alpha}_l$  (175) to be

$$\hat{\alpha}_l \equiv \alpha_l + \lambda |\alpha_l|. \quad (180)$$

Below is the code for the methods `mrqmin` and `mrqcof`.

```
void mrqmin(std::vector<std::string> param_names, vec &params,
    PDFBase& PDF_instance, Chi2 &chi2obj, mat &alpha, double &chisq,
```

<sup>5</sup>For comparison: The entire fitting process for an analytical function using the same algorithm takes a few milliseconds while each evaluation of  $\chi^2$  takes roughly 30 seconds for 1000 data points.

```

double &alamda){
/* One iteration of the Levenberg Marquard algorithm. Tries to reduce chi^2
of a fit by calculating a new set of trial parameters from the Hessian
matrix alpha and the vector of first derivatives beta.
Variables:
params: Current parameter values, set to new values if chisq decreases
alpha: Workspace during iterations, set to Hessian in final iteration
chisq: Lowest value of chisq found so far
alamda: Current value of lambda */
static double ochisq; // old chisq value
static int ma = params.size(); // Total number of parameters
static vec atry(ma), beta(ma), b_new(ma);
static mat a_new(ma, ma);
static vec old_diag(ma);

// Initial setup (called with lambda < 0)
if (alamda < 0.0) {
    alamda = 0.001; // Guess initial value
    // Initial calculation of alpha, beta and chi^2
    mrqcof(param_names, params, PDF_instance, chi2obj, chisq, alpha, beta);
    ochisq = chisq; // Save as old chi for next run
}

// Final run, calculates Hessian and returns before diagonal is changed
if (alamda == 0.0) {
    mrqcof(param_names, params, PDF_instance, chi2obj, chisq, alpha, b_new);
    return;
}

// Smooth variation between steepest descent and inverse Hessian method
// Large lambda corresponds to steepest descent (far from the minimum)
// Small lambda corresponds to inverse Hessian (close to minimum)
old_diag = alpha.diag(); // keep old values to reset in case of failure
alpha.diag() += alamda * abs(alpha.diag());
atry = params + solve(alpha, beta); // New trial parameters
// Check whether chisq improves with new params
mrqcof(param_names, atry, PDF_instance, chi2obj, chisq, a_new, b_new);
if (chisq < ochisq) { // If it improves, update and decrease lambda
    alamda *= 0.1;
    ochisq = chisq;
    alpha = a_new;
    beta = b_new;
    params = atry;
} else { // If it does not, reset everything and increase lambda
    alpha.diag() = old_diag;
    alamda *= 10.0;
}
}

void mrqcof(std::vector<std::string> param_names, vec params,
PDFBase& PDF_instance, Chi2 &chi2obj, double &chisq, mat &alpha, vec &beta){
/* Evaluates the Hessian matrix alpha and the gradient beta via finite
difference methods.
Variables:
alpha: Returns the Hessian for the current parameter values
beta: Returns the gradient for the current parameter values */
int ma = params.size();
vec fp(ma), fm(ma); // Chi2 with step forwards / backwards in each variable
double fpp, fmm; // Chi2 with step forwards / backwards in two variables
vec dp = abs(0.01*params)+0.0001; // Stepsize for the finite differences

```

```

// Chisq value for the current trial params
chisq = eval_chi(chi2obj, PDF_instance, param_names, params);

for (int i = 0; i < ma; i++){
    fp[i] = eval_chi(chi2obj, PDF_instance, param_names,
                    params+dp[i]*unit_vec(ma, i));
    fm[i] = eval_chi(chi2obj, PDF_instance, param_names,
                    params-dp[i]*unit_vec(ma, i));
}

for (int i = 0; i < ma; i++){
    for (int j = 0; j < i; j++){
        fpp = eval_chi(chi2obj, PDF_instance, param_names,
                        params+dp[i]*unit_vec(ma, i)+dp[j]*unit_vec(ma, j));
        fmm = eval_chi(chi2obj, PDF_instance, param_names,
                        params-dp[i]*unit_vec(ma, i)-dp[j]*unit_vec(ma, j));
        // alpha is defined as the second derivative times 0.5
        alpha.at(i,j) = alpha.at(j,i) =
            (fpp-fp[i]-fp[j]+2*chisq-fm[i]-fm[j]+fmm)/(4*dp[i]*dp[j]);
    }
    alpha.at(i,i) = (fp[i]-2*chisq+fm[i])/(2*dp[i]*dp[i]); // Diagonal
}
// beta is defined as first derivative times -0.5
beta = (-0.5)*(fp-fm) / (2*dp);
}

```

### 3.2.3 $K$ factors

The  $K$  factor is a handy tool to make calculations involving the computation of higher order theory predictions faster by approximating NLO (or higher order) results through LO computations. To do this, we first need to evaluate the desired observable ( $F_2$  in our case) both at LO and at NLO and then calculate the  $K$  factor defined as

$$K(x, Q; \mathbf{a}) = \frac{F_2^{\text{NLO}}(x, Q; \mathbf{a})}{F_2^{\text{LO}}(x, Q; \mathbf{a})}. \quad (181)$$

The NLO result for a different set of parameters  $\mathbf{a} + \delta\mathbf{a}$  is then approximated by

$$F_2^{\text{NLO}}(x, Q; \mathbf{a} + \delta\mathbf{a}) \approx K(x, Q; \mathbf{a})F_2^{\text{LO}}(x, Q; \mathbf{a} + \delta\mathbf{a}). \quad (182)$$

The approximation lies in the assumption that  $K(x, Q; \mathbf{a}) \approx K(x, Q; \mathbf{a} + \delta\mathbf{a})$  as long as  $\delta\mathbf{a}$  is sufficiently small. Whenever the parameters are changed by a substantial amount,  $K$  is calculated anew to keep a reasonable precision. To retain the best possible accuracy in the final result, we run the Levenberg-Marquardt algorithm from the initial parameter guess first with  $K$  factor and then the algorithm is restarted with the full theory starting from the parameters calculated with the  $K$  factor.



### 3.2.4 Examples

Now we want to test the general algorithm on a few test functions. First we start with a simple linear function of the form

$$f(x) = a_0 \cdot x + a_1. \quad (183)$$

Of course we will not be building an experiment where we expect linear behavior to try the method, but instead generate synthetic data by defining two parameters  $a_0^{\text{true}}, a_1^{\text{true}}$  and then evaluating the function over a domain of 40 values  $x_i \in [x_0, \dots, x_{N-1}]$ . Each point then gets a normally distributed random contribution with standard deviation  $\sigma$ . We choose  $a_0 = 0.4, a_1 = -1.5, \sigma = 0.025$ . Starting with parameters  $a_0 = -0.25, a_1 = 1.8$  the result is shown in figure 3.2. The best fit parameters lie close to the real ones, but a quantitative assertion of the fit

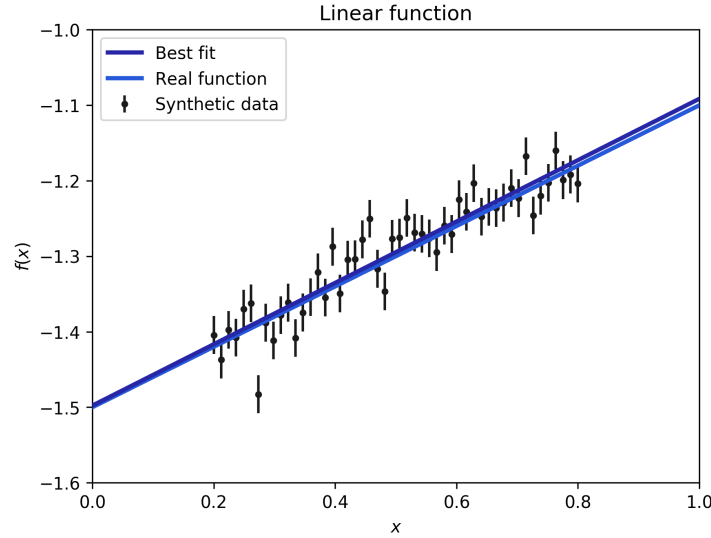


Figure 3.2: Best fit for the synthetic data ( $a_0 = 0.4064, a_1 = -1.4978$ ).

quality requires further tools, that will be introduced in chapter 3.3.

Next we want to try the algorithm on a harder problem, namely the sum of three Gaussians. The fitting of Gaussian distributions with the algorithm is very sensitive towards the initial parameters. Especially the position of the peaks has to be close to the real ones or else the algorithm does not converge. This is due to the fact, that it is easy to find small gaussian shaped bumps in the statistical noise which can lead the algorithm towards shallow local minima of  $\chi^2$ . Fitting 9 parameters also requires more data points to reach conclusive results, so the number of points is increased to 200. The model function is

$$f(x) = a_0 \cdot \exp\left(-\left(\frac{x - a_1}{a_2}\right)^2\right) + a_3 \cdot \exp\left(-\left(\frac{x - a_4}{a_5}\right)^2\right) + a_6 \cdot \exp\left(-\left(\frac{x - a_7}{a_8}\right)^2\right) + a_9, \quad (184)$$

with the "true" parameters  $\{0.7, 0.2, 0.05, 0.9, 0.7, 0.07, 0.5, 0.5, 0.2, 0.1\}$  and the initial parameters  $\{0.75, 0.25, 0.15, 0.5, 0.75, 0.15, 0.4, 0.4, 0.3\}$ . After 19 iterations the algorithm reaches the

best fit shown in figure 3.3. This time the generated data does not cover a large enough area

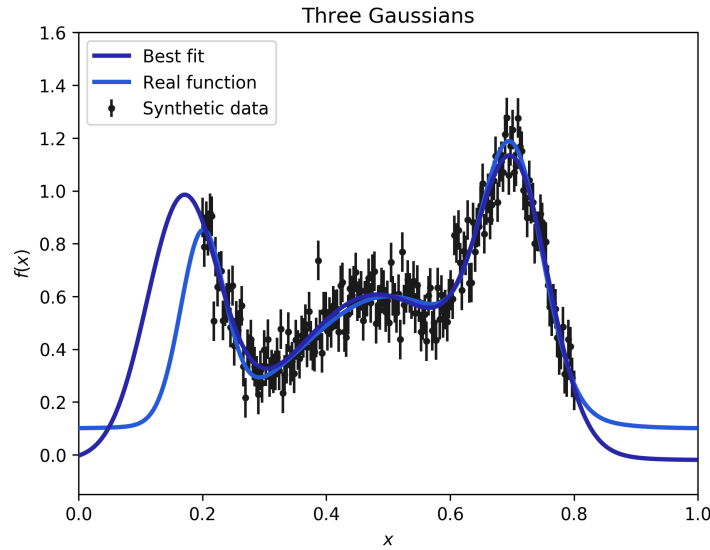


Figure 3.3: Best fit for the synthetic data. Parameters:  $\{0.9516, 0.1681, 0.0826, 0.9650, 0.7030, 0.0804, 0.6268, 0.4816, 0.1993, -0.0199\}$ .

to accurately determine all parameters of the real function. Especially the leftmost peak and the offset differ from the real function because of the lacking data in those areas. Across the interval  $[0.2, 0.8]$ , where data exists the fit agrees very well with the true function, such that the total  $\chi^2$  still reaches a reasonable value of 216.024 ( $\frac{\chi^2}{\nu} \approx 1.14$ ). We expect the uncertainties of the fit to reflect these facts, but discussion of the uncertainty of the fit is again relegated to the chapter 3.3.

As the last example we want to take a look at a case where the model function is different from the function used to generate the data. We want to use a polynomial of sixth order  $f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + a_6x^6$  to describe data that is again generated from the sum of three Gaussians. The results are shown in figure 3.4.

Unsurprisingly the fit is worse than before with  $\chi^2 = 726.3$  or  $\frac{\chi^2}{\nu} \approx 3.76$  which tells us that a polynomial of this degree is not sufficiently flexible to model the data. It does, however still capture the rough shape of the data. The biggest disagreement between real function and model can be seen in the left peak, where the polynomial fails to accurately replicate the Gaussian at  $x = 0.1$ . If the model would be closer to the left peak in its shape, then it would descend even steeper for  $x \rightarrow 0$  and thereby miss the points where the real function starts flattening, which gives this shape an overall lower  $\chi^2$ . Outside of the data region the polynomial model quickly diverges from the real function due to the simple fact that the convergence towards a finite value cannot be replicated by a polynomial of finite degree.

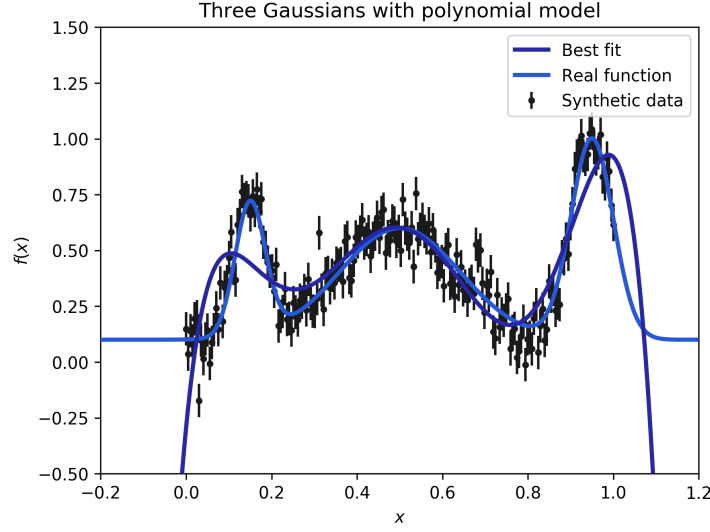


Figure 3.4: Best fit for the synthetic data. Parameters:  $\{-0.293, 19.79, -176.2, 674.3, -1226, 1045, -335.3\}$ .

### 3.3 The Hessian formalism for uncertainties

Now that we have a method to fit the PDFs we need to quantify our uncertainty of the PDFs themselves and that of predictions made for quantities dependent on the PDFs. This is difficult for many reasons. On the one hand there are errors from theory e.g. higher-order QCD corrections, nuclear effects and power suppressed corrections. On the other hand the errors from published data sets are often only provided in uncorrelated form with a single overall normalization uncertainty. Additionally, published errors for some experiments seem to fail standard statistical tests e.g. having  $\chi^2$  differ significantly from 1 per degree of freedom, which makes these data sets rather "improbable". Lastly some individual data sets seem to require mutually incompatible PDF parameters. This chapter is devoted to explaining how the uncertainty of the PDFs can be assessed systematically using the Hessian method from [18].

#### 3.3.1 Parameterization of PDFs and the effective $\chi^2$

The Hessian approach is based on a quadratic approximation of  $\chi^2_{\text{global}}$  near its minimum, which defines the best fit. Another possible approach is based on the Lagrange multiplier method [35]. This has the advantage of not requiring any approximations at the cost of not providing complete information about the minimum's neighbourhood. In this thesis we will focus on the Hessian approach.

The theory contains a set  $\{a_i\} = \{a_1, \dots, a_d\}$  of  $d$  free parameters that characterize the non-perturbative input. Fitting the theory to the data determines these  $\{a_i\}$  and thereby the PDFs. The uncertainty of the fit due to both experimental and theoretical errors is determined from the range of permissible  $\Delta\chi^2$  for the fit. The specific forms are not crucial, as long as the parameterization is flexible enough to model the behaviour of the true PDFs to sufficient accuracy. Too much flexibility is problematic however, since it would lead to some parameters

being poorly determined at the minimum of  $\chi^2$  if the amount or quality of experimental data is not high enough.

### Effective $\chi^2$ function

Every PDF set  $S$  is defined through a parameter set  $\{a_i(S)\}$  and an effective global  $\chi_{\text{global}}^2(S)$  function as defined in (160), that gives us a measure of how well the data is fit by the theory with this parameter set.

With this effective  $\chi^2$  function we define the best estimate of the true PDF as the parameter set  $S_0$ , where  $\{a_i(S_0)\}$  minimizes  $\chi_{\text{global}}^2$ . Information about the uncertainties is gained by exploring the neighbourhood of the minimum in parameter space. When the parameters are changed from the minimum  $\chi_{\text{global}}^2$  increases by an amount  $\Delta\chi_{\text{global}}^2$ . We can then define the relevant neighbourhood of "acceptable fits" through a tolerance parameter  $T$ :

$$\Delta\chi_{\text{global}}^2 \leq T^2 \quad (185)$$

The Hessian method will provide a method of calculating the variation of the predictions of the PDFs in this neighbourhood, as long as the quadratic approximation of  $\chi_{\text{global}}^2$  is sufficiently accurate inside the entire range of  $T$ . Due to deviations from ideal statistical expectations in some data sets and signs of incompatibility between sets,  $T$  must be chosen significantly higher than it would be in a case with ideal statistics. The specifics of estimating  $T$  will be discussed later on in section 3.3.3.

### 3.3.2 The Hessian formalism

The Hessian method is based on a quadratic expansion of the  $\chi^2$  function around its minimum. The Hessian matrix is the matrix of second derivatives of  $\chi^2$  at the minimum. The eigenvectors of the Hessian are used to produce an eigenvector basis set of PDFs from which uncertainties of physical observables can be calculated. The general concept of the approach can be explained with figure 3.5, which uses just two parameters for simplicity. Every PDF set  $S$  corresponds to a point in the  $d$ -dimensional parameter space, specified either by the coordinates  $\{a_i(S)\}$  in the original parameter basis or by  $\{z_k(S)\}$  in the eigenvector basis. The Hessian matrix is diagonalized and the eigenvectors are rescaled by an iterative procedure so that the contours of constant  $\chi^2$  become spherical in the new basis. The scaling factor  $s_k$  depends on the corresponding eigenvalue and will give us information about the uncertainty in each direction.

#### Quadratic approximation of the Hessian matrix

We start by expanding  $\chi_{\text{global}}^2(S)$  as a Taylor series around its minimum at  $S_0$ , up to leading order. Dropping the *global* subscript and introducing  $\chi_0^2 = \chi^2(S_0)$ ,  $\{a_j^0\} = \{a_j(S_0)\}$ ,  $\{a_j\} =$

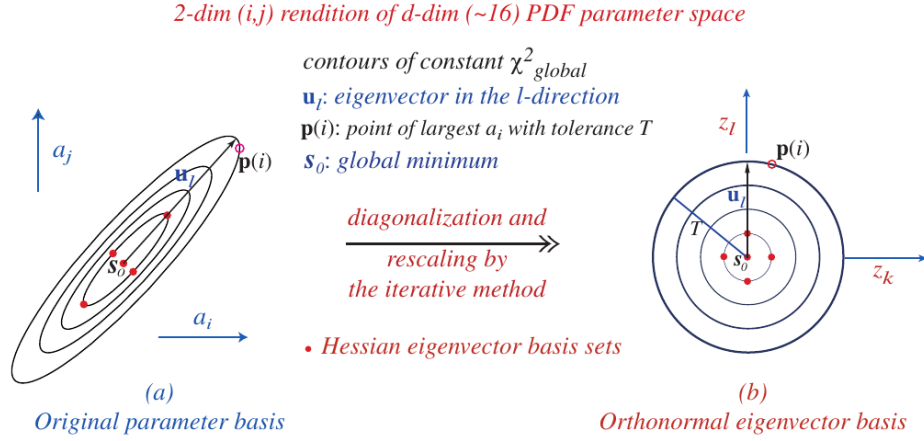


Figure 3.5: Illustration of the basic idea behind the formalism. Figure taken from [18].

$\{a_j(S)\}$  and the displacement from the minimum  $y_i = a_i - a_i^0$ , we get:

$$\Delta\chi^2 = \chi^2 - \chi_0^2 = \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d H_{ij} y_i y_j, \quad (186)$$

$$\text{with } H_{ij} = \left( \frac{\partial^2 \chi^2}{\partial y_i \partial y_j} \right) \Big|_{a_i = a_i^0}. \quad (187)$$

From now on we will also leave the  $(S)$  argument implicit unless needed for clarity. A complete set of orthonormal eigenvectors  $v_{ik}$  of the  $H_{ij}$  can be defined by:

$$\sum_{j=1}^d H_{ij} v_{jk} = \epsilon_k v_{ik} \quad \text{and} \quad \sum_{i=1}^d v_{il} v_{ik} = \delta_{lk}, \quad (188)$$

where  $\{\epsilon_k\}$  are the eigenvalues and  $\delta_{lk}$  the unit matrix. The displacements from the minimum can be expressed in terms of the eigenvectors as

$$y_i = \sum_{k=1}^d v_{ik} s_k z_k, \quad (189)$$

where  $s_k$  are scale factors normalizing the new parameters  $z_k$  such that

$$\Delta\chi^2 = \sum_{k=1}^d z_k^2 \leq T^2, \quad (190)$$

which makes the relevant neighborhood a hypersphere of radius  $T$ . The necessary scale factors  $s_k$  are approximately  $\sqrt{\frac{2}{\epsilon_k}}$ , but some adjustments are made to improve the quality of the quadratic approximation.

### Eigenvalues of the Hessian matrix and PDF eigenvector basis sets $S_l^\pm$

Squaring and then summing over  $i$  in (189) and using (188) gives us the square of the distance

in parameter space from the minimum  $\chi_0^2$ :

$$\sum_{i=1}^d (a_i - a_i^0)^2 = \sum_{k=1}^d (s_k z_k)^2. \quad (191)$$

Since  $s_k \approx \sqrt{\frac{2}{\epsilon_k}}$ , an eigenvector with a large eigenvalue therefore corresponds to a direction in  $\{a_i\}$  space where  $\chi^2$  rises rapidly. This means that parameters are tightly constrained by the given data. Accordingly, a small eigenvector allows larger deviations due to (190).

The orthogonal matrix transforming between the original parameter basis and the eigenvector basis is given by  $v_{ik}$ . Since the normalized coordinates  $\{z_k\}$  are more convenient to work with, we define a matrix

$$M_{ik} \equiv v_{ik} s_k, \quad (192)$$

such that

$$a_i - a_i^0 = \sum_{k=1}^d M_{ik} z_k. \quad (193)$$

This matrix holds information about the physics of the global fit and can also be used to search for possible improvements of the parameterization. The eigenvectors provide an orthonormal basis in the PDF parameter space, which gives us a simple parameterization of the parton distributions near the global minimum  $S_0$ .

These eigenvector basis sets are defined by displacements of a standard magnitude  $t$  along one of the eigenvector directions:

$$z_k(S_l^\pm) = \pm t \delta_{kl}. \quad (194)$$

As an explicit example,  $S_1^+$  is defined by  $(z_1, \dots, z_d) = (t, 0, \dots, 0)$ . We choose  $t$  to be equal to  $\frac{T}{\sqrt{N_{\text{params}}}}$  as will be explained later on. In practice the specific value of  $t$  does not have much influence on the final uncertainties as long as it lies in a reasonable region.

The parameter set  $\{a_i\}$  that specifies the eigenvector basis  $S_l^\pm$  is given by

$$a_i(S_l^\pm) - a_i^0 = \pm t M_{il} \quad (195)$$

$$\implies a_i(S_l^+) - a_i(S_l^-) = 2t M_{il}. \quad (196)$$

This can be interpreted as a difference equation such that the matrix element  $M_{il}$  of the transformation matrix is equal to the gradient of parameter  $a_i$  along the direction of  $z_l$ .

### Calculating uncertainties using the eigenvector basis sets $\{S_l^\pm\}$

Consider a variable  $X(S)$  that depends on the PDFs. It could be a physical quantity such as a certain processes cross section, a particular PDF at specific  $x$  and  $Q$  or even one of the parameters  $a_i$ . The best fit estimate for  $X$  is then  $X_0 = X(S_0)$ . To find the uncertainty, all one has to do is evaluating  $X$  for each of the  $2d$  eigenvector basis sets  $S_l^\pm$ . In the linear approximation of the Hessian method, the gradient of  $X$  in the  $z$  representation can then be

calculated as

$$\frac{\partial X}{\partial z_k} \approx \frac{X(S_k^+) - X(S_k^-)}{2t}. \quad (197)$$

It is useful to define

$$D_k(X) = X(S_k^+) - X(S_k^-), \quad (198)$$

$$D(X) = \sqrt{\sum_{k=1}^d (D_k(X))^2}, \quad (199)$$

$$\hat{D}_k(X) = \frac{D_k(X)}{D(X)}, \quad (200)$$

so that  $D_k(X)$  is a vector in the gradient direction and  $\hat{D}_k$  is the unit vector in that direction. The gradient direction is the direction in which  $X$  varies the fastest. Therefore the largest variation in  $X$  allowed by (190) are obtained by displacing  $S_0$  along the gradient direction by a distance  $\pm T$ . This gives us the uncertainty

$$\Delta X = \sum_{k=1}^d (T \hat{D}_k) \frac{\partial X}{\partial z_k}. \quad (201)$$

Using (197) - (200) the uncertainty can be written in the more convenient form

$$\Delta X = \frac{T}{2t} D(X). \quad (202)$$

Often it is also important to construct the PDF sets  $S_X^+$  and  $S_X^-$  that create the extreme values  $X = X_0 \pm \Delta X$ . Their coordinates in  $z$  space are

$$z_k(S_X^\pm) = \pm T \hat{D}_k(X). \quad (203)$$

Using (193) and (196) the physical parameters that give the extremal values for the variable  $X$  are given by:

$$a_i(S_X^\pm) - a_i^0 = \frac{\pm T}{2t} \sum_{k=1}^d \hat{D}_k(X) [a_i(S_k^+) - a_i(S_k^-)]. \quad (204)$$

We can also estimate the uncertainties of each individual parameter. With  $X = a_i$  we get from (196) and (198):

$$D_k(a_i) = a_i(S_k^+) - a_i(S_k^-) = 2t M_{ik}. \quad (205)$$

Plugging this into (202) gives

$$\Delta a_i = T \sqrt{\sum_k M_{ik}^2}. \quad (206)$$

Now the only thing left to do is finding the values of the two parameters  $T$  and  $t$ .

### 3.3.3 Estimation of $T$ and $t$

To estimate the parameters  $T$  and  $t$  we need to take a look at the properties of the  $\chi^2$ -distribution. Its probability density function is

$$P(\chi_\nu^2; \nu) = \frac{1}{2^{\nu/2} \Gamma(\nu/2)} (\chi_\nu^2)^{\nu/2-1} e^{-(\chi_\nu^2)/2}, \quad (207)$$

where  $\nu$  is the number of degrees of freedom. For large  $k \gg 1$ , the distribution has a mean of  $\langle \chi_\nu^2 \rangle = \nu$  and a standard deviation of  $\sigma = \sqrt{2\nu}$ . In the case of perfect statistics we can use  $T^2 = \Delta \chi^2 = \sigma$ , but we need to account for the fact that the actual experimental errors do not follow perfect normal distributions. To do this we rescale  $T^2$  by a factor  $\frac{\chi_{\text{fit}}^2}{\langle \chi_\nu^2 \rangle} \approx \frac{\chi_{\text{fit}}^2}{\nu}$  giving us overall

$$T = \sqrt{\frac{\chi_{\text{fit}}^2}{\nu} \sqrt{2\nu}}. \quad (208)$$

The parameter  $t$  is supposed to be the standard magnitude of the displacement along one eigenvector direction and should therefore be somewhat smaller than  $T$ . We can get its value from (202) by using  $\chi^2$  as the observable. In that case we get

$$T^2 = \frac{T}{2t} \sqrt{\sum_i^{N_{\text{params}}} (2t^2)^2} \quad (209)$$

$$T^2 = \frac{T}{2t} \sqrt{N_{\text{params}}} 2t^2 \quad (210)$$

$$\implies t = \frac{T}{\sqrt{N_{\text{params}}}}. \quad (211)$$

To make sure uncertainties are not underestimated  $t$  is sometimes chosen to be somewhat higher than (211) suggests, but in our case trials have shown that increases in  $t$  produce no noticeable difference in the resulting uncertainties.

### 3.3.4 Implementation in C++

Using the linear algebra capabilities of the armadillo C++ library the calculation of the transformation matrix  $M$  is very straightforward. We use the `eig_sym` function to calculate the eigenvalues and -vectors of the symmetric hessian and then calculate the scale factors as  $s_k = \sqrt{\frac{2}{\epsilon_k}}$ , where  $\epsilon_k$  are the eigenvalues. Each column of the matrix is then calculated as the eigenvector times the corresponding scale factor.



```

mat get_M_trans(mat &hessian){
  /* Calculates the transformation matrix from the Hessian's eigenvectors and
     uses the eigenvalues for approximate normalization.
     Variables:
     hessian: Hessian matrix */
  vec eigval;
  mat eigvec;
  mat M_trans(size(hessian));

  // Calculate eigenvalues and eigenvectors of (symmetric) Hessian
  eig_sym(eigval, eigvec, hessian);
  // Calculate scale factors from eigenvalues.
  vec scale = sqrt(2.0*ones<vecfor (int i = 0; i<eigval.size(); i++){
    M_trans.col(i) = eigvec.col(i)*scale[i];
  }
  return M_trans;
}

```

The scaling of the columns of  $M$  is then adjusted, so that

$$\Delta\chi^2(\mathbf{a} \pm t \cdot M_i) \equiv \chi^2(\mathbf{a}_0 + z \cdot M_i) - \chi^2(\mathbf{a}_0) = t^2 \quad (212)$$

for each column  $M_i$  of the matrix  $M$ . Since  $\Delta\chi^2$  is not shaped perfectly parabolic, it is common to use

$$\Delta\chi^2(\mathbf{a}_0 + t \cdot M_i) + \Delta\chi^2(\mathbf{a}_0 - t \cdot M_i) = 2t^2, \quad (213)$$

so that the average of the two sides is scaled properly. This assumes that the underestimation of the uncertainties on one side and overestimation on the other side cancel out and hence the result is close to the real uncertainty. We use

$$\min(\Delta\chi^2(\mathbf{a}_0 + t \cdot M_i), \Delta\chi^2(\mathbf{a}_0 - t \cdot M_i)) = t^2 \quad (214)$$

instead to make sure the uncertainties are always overestimated on one side and exact on the other. The rescaling has to be done iteratively because the exact form of the  $\Delta\chi^2$  curves is not known. In theory, it should be optimal to rescale  $M$  by a factor of  $(t \cdot \min(\Delta\chi^2(\mathbf{a}_0 + t \cdot M_i), \Delta\chi^2(\mathbf{a}_0 - t \cdot M_i))^{-1/2})$  in each iteration, but in practice it is faster to use the square root of that factor to avoid overshooting the desired scale, which can lead to oscillations between scaling up and down. This procedure is implemented in the function `get_M_improved`.

```

mat improve_M(std::vector<std::string> param_names, vec &params, double t,
  PDFBase& PDF_instance, Chi2 &chi2obj, mat &hessian, double &chisq){
  /* Returns the transformation matrix between the normalized eigenvector
     basis and the original basis. Assumes that Delta chisq is convex
     along each EV direction and that the minimum of chisq along each EV lies
     somewhat close to the middle of the interval [-t,t]. Ensures that errors
     are not underestimated, if Delta chisq is not symmetrical.
     Variables:
     t: Standard magnitude of displacement along each eigenvector direction
     hessian: Hessian matrix */
  int mfit = params.size();

```

```

vec z = linspace<vec>(-t,t,2); // z = +/- t
mat M_trans(mfit, mfit); // Transformation matrix
M_trans = get_M_trans(hessian); // Matrix with approximate scaling

for (int i = 0; i < mfit; i++){ // Iterate over all EV directions
    double rescale = 0;
    // Repeat until most recent change was less than 1%
    while(rescale > 1.01 || rescale < 0.99){
        // Rescale such that the smaller Delta chisq is equal to t^2
        rescale = t*t/min(dchi_single(param_names, params, i, PDF_instance,
                                     chi2obj, M_trans, chisq, z));
        // Using a single sqrt would be faster but risks overshooting
        M_trans.col(i) = M_trans.col(i)*sqrt(sqrt(rescale));
    }
}
return M_trans;
}

```

This function makes use of a helper function `dchi_single` to calculate the values of  $\Delta\chi^2(\mathbf{a} \pm t \cdot M_i)$ . In general `dchi_single` returns the values of  $\Delta\chi^2(\mathbf{a} + z \cdot M_i)$  for all values  $z$  contained in the vector  $z$ . This function is also used to obtain the  $\Delta\chi^2$  curves as seen for example in Figure 4.8.

```

vec dchi_single(std::vector<std::string> param_names, vec &params, int dir,
PDFBase& PDF_instance, Chi2 &chi2obj, mat &M_trans, double &chisq, vec z){
/* Calculates the deviation of chisq along the given eigenvector direction for
all displacement in z.
Variables:
dir: index of the EV direction
M_trans: Transformation matrix between original and EV basis
z: Displacements where Delta chisq is evaluated */
int mfit = params.size(); // Number of parameters
int N_t = z.size(); // Number of displacements
vec dchi(N_t); // Return vector with all Delta chisq values

// Iterate over all displacements and calculate Delta chisq for each
for (int j = 0; j < N_t; j++){
    dchi[j] = eval_chi(chi2obj, PDF_instance, param_names,
                      params + z[j]*M_trans.col(dir))-chisq;
}
return dchi;
}

```

Finally we need a method that gives us the sets of eigenvector parameters. The function `get_EV_params` returns a matrix where each column corresponds to a set of eigenvector parameters. The zeroth column of the returned matrix is just the best-fit parameter set  $\mathbf{a}$ . The columns  $2n+1$  and  $2n+2$  with  $n = 0, \dots, N_{\text{params}}$  are then filled with  $(\mathbf{a} + t \cdot M_n)$  and  $(\mathbf{a} - t \cdot M_n)$  respectively.

```

mat get_EV_params(vec params, mat M_trans, double t){
/* Takes the best fit parameters, transformation matrix and the displacement t
to calculate all eigenvector parameter sets.
Variables:
params: best fit parameters
M_trans: Transformation matrix between original and eigenvector basis
t: Standard magnitude of displacement along each eigenvector direction */

```

```

int mfit = params.size();
mat ev_params(mfit, 2*mfit+1);
ev_params.col(0) = params; // 0th EV is the best fit

// Calculate eigenvector parameter sets
for (int i = 0; i<2*mfit; i++){
    vec dp = t*M_trans.col(i/2); // Displacements along i-th eigenvector
    ev_params.col(i+1) = params + dp; // Positive direction
    i++;
    ev_params.col(i+1) = params - dp; // Negative direction
}
return ev_params;
}

```

### 3.3.5 Fit improvement

Because we need to use finite difference approximations in the fitting procedure, the algorithm might fail to find further improvements of  $\chi^2$  when it is already close to the minimum. However, the eigenvector directions obtained from the Hessian provide a natural basis to explore the  $\chi^2$  space close to a minimum. Therefore we can use them to scan the parameter space along the eigenvectors and check whether the Levenberg-Marquardt algorithm actually provided an accurate minimum and apply corrections to the parameters if necessary. We use the matrix  $M$  before rescaling because the calculation of the improved matrix assumes that the minimum along each direction already lies close to the center of the interval  $[-t, t]$ .

At the beginning of the algorithm the Hessian and the corresponding transformation matrix  $M$  are calculated. Then we iterate over all eigenvectors scanning the  $\chi^2$  space along them and setting the parameters to the value where a local minimum is found. If the parameters are found to be at the minimum along a direction  $i$  already we save that fact by setting a variable `centered[i]` to 1, but whenever a new minimum is found along direction  $i$ , all values `centered[j]` for  $j \neq i$  are reset to 0. This procedure is repeated until `centered[i]=1` for all  $i$ . The C++ implementation of this algorithm is given below:

```

void improve_fit(std::vector<std::string> param_names, vec &params, double t,
    PDFBase& PDF_instance, Chi2 &chi2obj, mat &hessian, double &chisq){
/* Improve the fit by looking at the dchi curves along the eigenvectors
and setting the parameters to their minima.
Variables:
t: Maximum displacement along each eigenvector
hessian: Hessian matrix */
int mfit = params.size(); // Number of parameters
int N_t = 13; // Number of points where dchi is evaluated
int mid = N_t/2;
uvec centered = zeros<uvec>(mfit); // Saves which curves are centered
vec z = linspace<vec>(-t,t,13); // Values of z where dchi is evaluated
vec z3 = linspace<vec>(-t/mid, t/mid, 3);
vec dchi(N_t); // Delta chisq values for each z
vec b(mfit); // Dummy variable needed by mrqcof
double chi_left, chi_right; // Delta chisq at the left- and rightmost z value

while(!all(centered)){ // Iterate until all curves are centered at z=0
    mat M_trans = get_M_trans(hessian); // Calculate M with current Hessian
    for (int i = 0; i<mfit ;i++){ // Iterate over all EV directions

```

```

if(!centered[i]){ // Skip EV if it is known to be centered
// Compare z=0 with the smallest displacement in +/- direction
if(all(dchi_single(param_names, params, i, PDF_instance, chi2obj,
                    M_trans, chisq, z3) > -0.01)){
    centered[i] = 1;
    continue; // Skip parameter if it is centered
}
// Calculate full dchi curve if EV is not centered yet
dchi = dchi_single(param_names, params, i, PDF_instance, chi2obj,
                    M_trans, chisq, z);
// Update parameters to the minimum found on the curve
params = params + z[index_min(dchi)]*M_trans.col(i);
chisq += dchi[index_min(dchi)]; // Update chisq
// If minimum was outside the z interval check the same EV again
if(index_min(dchi)==0 || index_min(dchi)==N_t-1){
    i--;
    continue;
} else {
    // All other directions are no longer known to be centered
    centered = zeros<uvec>(mfit);
    centered[i] = 1; // Save this direction as centered
}
}
}
// Recalculate Hessian after changing parameters
mrqcof(param_names, params, PDF_instance, chi2obj, chisq, hessian, b);
}
}

```

### 3.3.6 Examples

In section 3.2.4 we postponed the discussion of the uncertainties of the two example fits for the lack of the necessary tools. The Hessian formalism now allows us to rectify this. Additionally, the Hessian formalism works very well with the Levenberg-Marquardt algorithm because the algorithm already gives us the Hessian for free when we fit a model function. Going back to the linear model we can estimate  $T$  to be 3.2, because  $T^2 = \Delta\chi^2 \approx \sqrt{2\nu} = \sqrt{2(50-2)} \approx 10$  and accordingly we get  $t = \frac{T}{\sqrt{N_{\text{params}}}} \approx 2.3$ .

The band of uncertainty already lets us see one of the nice features of the Hessian method: the band of uncertainties is smallest around  $x \approx 0.5$ . If the method would only yield the uncertainty of each individual parameter as a result, the band of uncertainties would always be thinnest at  $x = 0$  because the contribution of  $a_0$  (and therefore its uncertainty) vanishes there. Looking at the eigenvector curve sets however we see the red coloured one mainly affects the offset along the  $y$ -axis while the green ones are what leads to the band of uncertainty being the smallest around  $x \approx 0.5$ . The first set of eigenvectors changes the parameters by  $\Delta a_0^{(\pm)} = \mp 0.04, \Delta a_1^{(\pm)} = \pm 0.02055$ , because this set combines high offset with flat slope and low offset with steep slope, it is mostly responsible for the symmetric shape of the error band around the middle of the data set. The two curves intersect at  $x = 0.51375$  meaning this set of eigenvector curves gives no contribution to the error band at that point. The second eigenvector set changes the parameters by  $\Delta a_0^{(\pm)} = \pm 0.003, \Delta a_1^{(\pm)} = \pm 0.00565$ . Even though the changes to both parameters are significantly smaller than they were for the first set, they still lead to

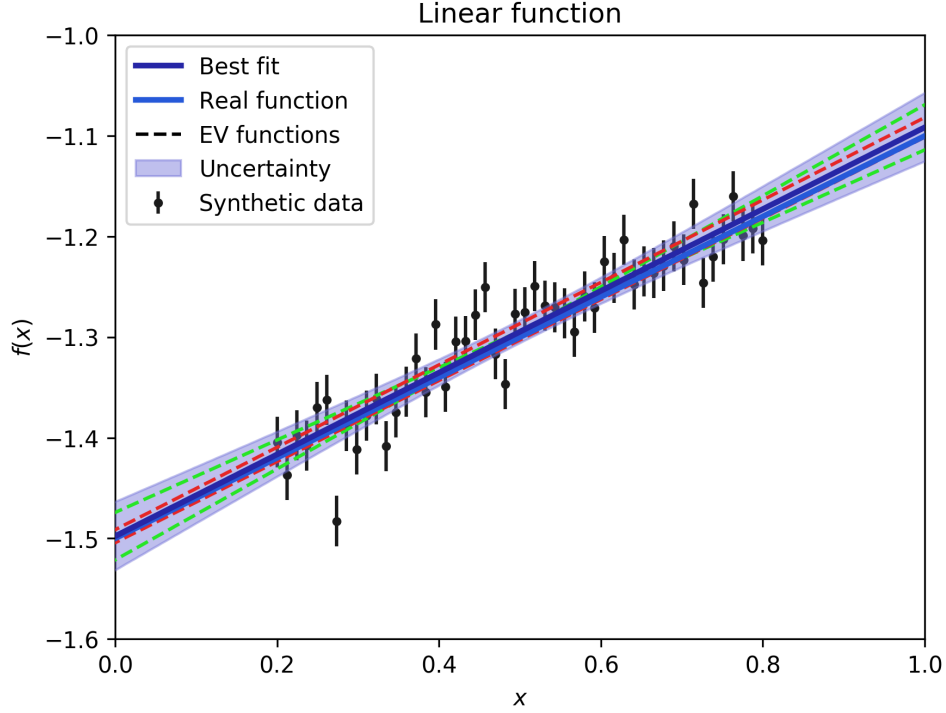


Figure 3.6: Best fit for the linear model with uncertainties and the two sets of functions with parameters shifted along the eigenvector directions. The two functions belonging to the same eigenvector share the same colour.

the same change in  $\chi^2$  because in this case higher offset is combined with steeper slope.

Next we take a look at the uncertainties of the fit for the three Gaussians, where the left-most peak of the function lies outside of the range covered by the synthetic data. With  $N_{\text{data}} = 200, N_{\text{params}} = 10$  and  $\chi^2_{\text{fit}} = 216$  we get  $T = 4.7, t = 1.6$ . The resulting uncertainties and eigenvector curves are shown in figure 3.7. Across the region populated with data the uncertainties are expectedly very small. As soon as the data points stop the uncertainties quickly blow up on both sides. The reason for that is that the data stops before the shape of the function has been completely determined. The offset  $a_9 = -0.0199 \pm 0.4821$  is not clearly discernible on either side of the curve because the region filled with data is too narrow, which causes large uncertainties for  $x < 0.2$  and  $x > 0.8$ . The uncertainty for  $x < 0.2$  is further increased by the fact that the peak of the left Gaussian lies just on the edge of the data causing it to be ill determined. Even though the offset comes into play for every value of  $x$  equally, the uncertainty of the overall function can be significantly smaller than the uncertainty of the offset itself. The reason for that is that the eigenvector directions responsible for the large deviations in  $a_9$  affect the other parameters in such a way that the value of  $f(x)$  remains mostly unchanged across the interval  $[0.2, 0.8]$ . It can also be seen from the colours that the same two eigenvector curve sets responsible for the bulk of the uncertainty of the offset are responsible for the uncertainty of the left peak's shape. Since there is not enough data to tightly constrain the shape of this peak the eigenvectors cannot negate the effect of offset in this region, like they do in the region filled with data.

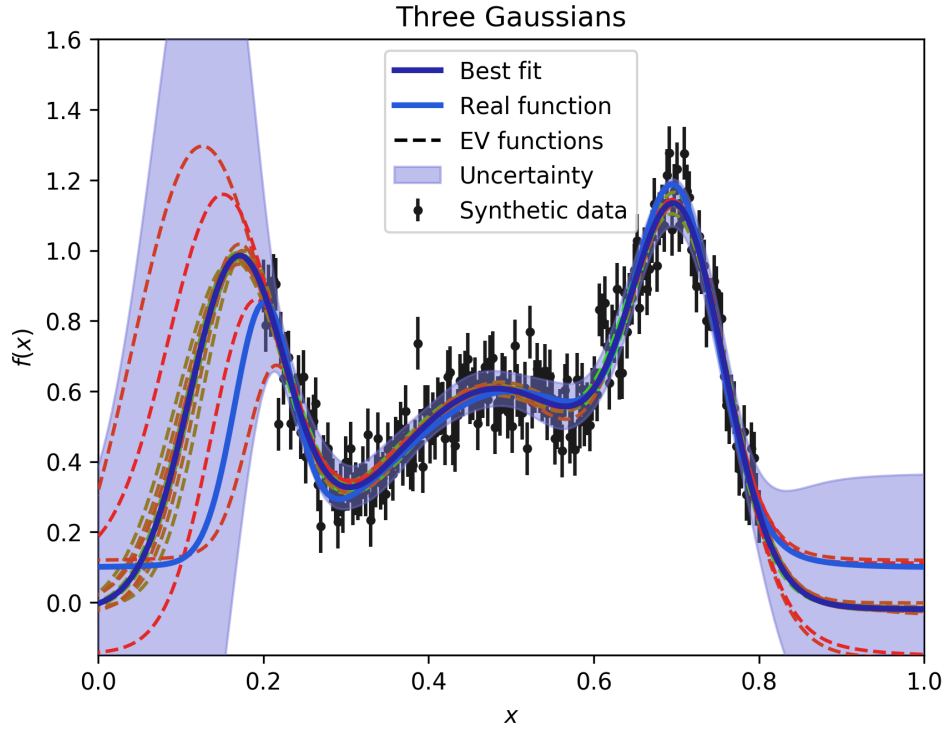


Figure 3.7: Best fit for the three Gaussian model with uncertainties and the 9 sets of functions with parameters shifted along the eigenvector directions. The two functions belonging to the same eigenvector share the same colour.

Lastly, we take another look at the fit of the three Gaussians with a polynomial model. The large value of  $\chi^2_{\text{fit}} = 726.3$  gives us  $T = 8.8, t = 3.3$ . While the fit does a decent job at replicating the rough shape of the data it misses the shape of the left peak. The real function and most of the data points are relatively close to the band of uncertainty, but on the left side it seems that the uncertainty fails to capture the disagreement between fit and data. Perhaps unintuitively, large deviation between data and fit does not imply large uncertainty. The uncertainty is defined by the flexibility of the model in the region around the minimum of  $\chi^2$  where  $\Delta\chi^2 < T^2$ . This means that data far away from the model can still tightly constrain the parameters. However, this can only happen if the model is too inflexible to accurately describe the data. Therefore a tight band of uncertainty in a region where the data disagrees with the fit is a sign of an insufficient model.

One further noteworthy fact about the Hessian method, that cannot be seen in any of the three examples shown, is that not all eigenvector functions always lie inside of the band of uncertainties. This seems unintuitive at first, but it happens because symmetric variation of the parameters around their best estimate does not necessarily lead to eigenvector curves that always lie on opposite sides of the best fit curve. For an eigenvector curve to lie outside of the error band its partner curve has to lie on the same side of the best fit function and the distance between the two has to be smaller than the width of the uncertainty.

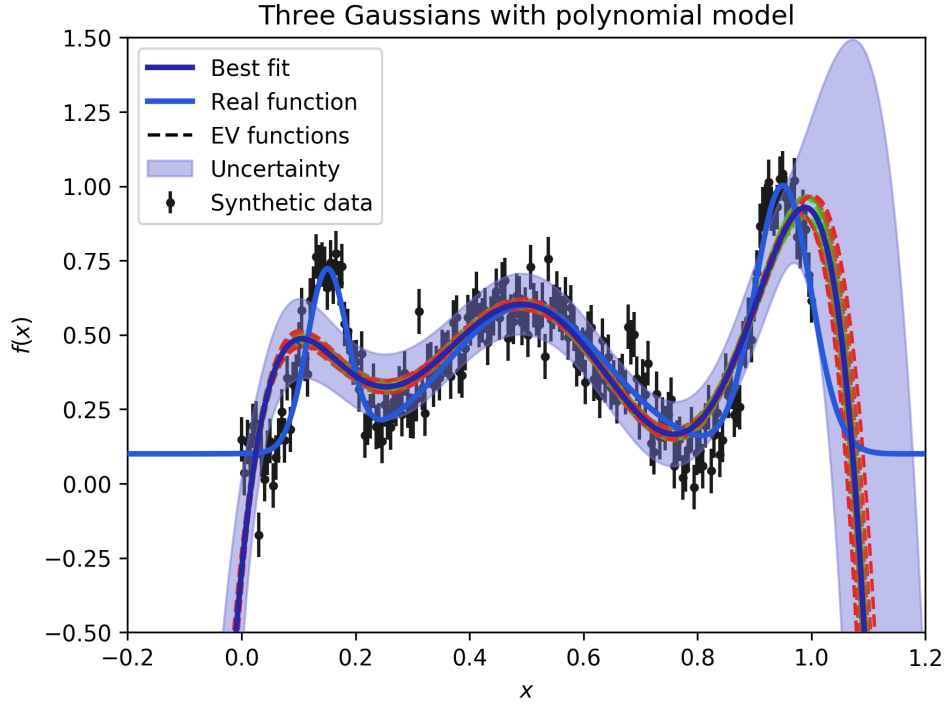


Figure 3.8: Best fit for the polynomial model of the three Gaussians with uncertainties and the 7 sets of functions with parameters shifted along the eigenvector directions. The two functions belonging to the same eigenvector share the same colour.

## 4 Data analysis

This section will cover the preparation of the data that will later be used to fit PDFs and other things. In general we use the DIS data for proton and deuteron  $F_2$  from the BCDMS [20], [21] and NMC [19] experiments.

### 4.1 Preparation of data

We discard any data where only one target is available at a certain point in the kinematic  $(x, Q^2)$ -plane and apply kinematic cuts for low  $Q^2$  and  $W^2$ . We cut  $Q^2$  below  $4 \text{ GeV}^2$  to avoid the necessity of higher twist corrections and  $W^2$  below  $20.25 \text{ GeV}^2$  to ensure that the process is actually deep inelastic scattering instead of elastic scattering ( $W = m_{\text{Proton}}$ ) or excitation of Delta resonances. Figure 4.1 shows the kinematic region of the available data and the applied cuts. We see that the BCDMS data is unaffected by the cuts and that the  $W^2$  cut does not constrain the given data at all beyond what has already been done by the  $Q^2$  cut. We are left with a total of 992 data points.

As a first trial, we will compare the data with theoretical predictions for  $F_2$  based on the CTEQ6.6 PDFs [41]. The uncertainties are computed based on the Hessian formalism introduced previously. This is very convenient in this case because the eigenvector PDF sets can be downloaded along with the PDFs from the CTEQ website<sup>6</sup>.

<sup>6</sup><https://hep.pa.msu.edu/cteq/public/cteq6.html>



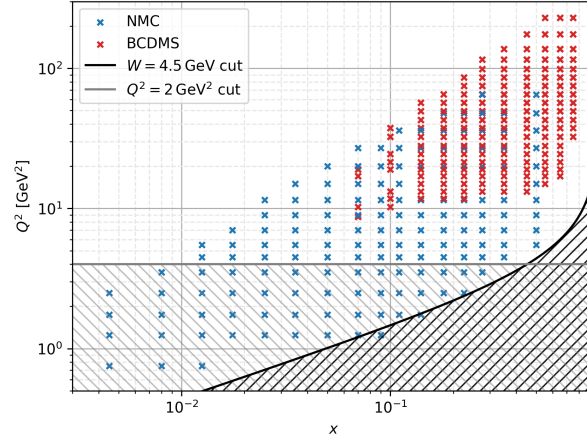
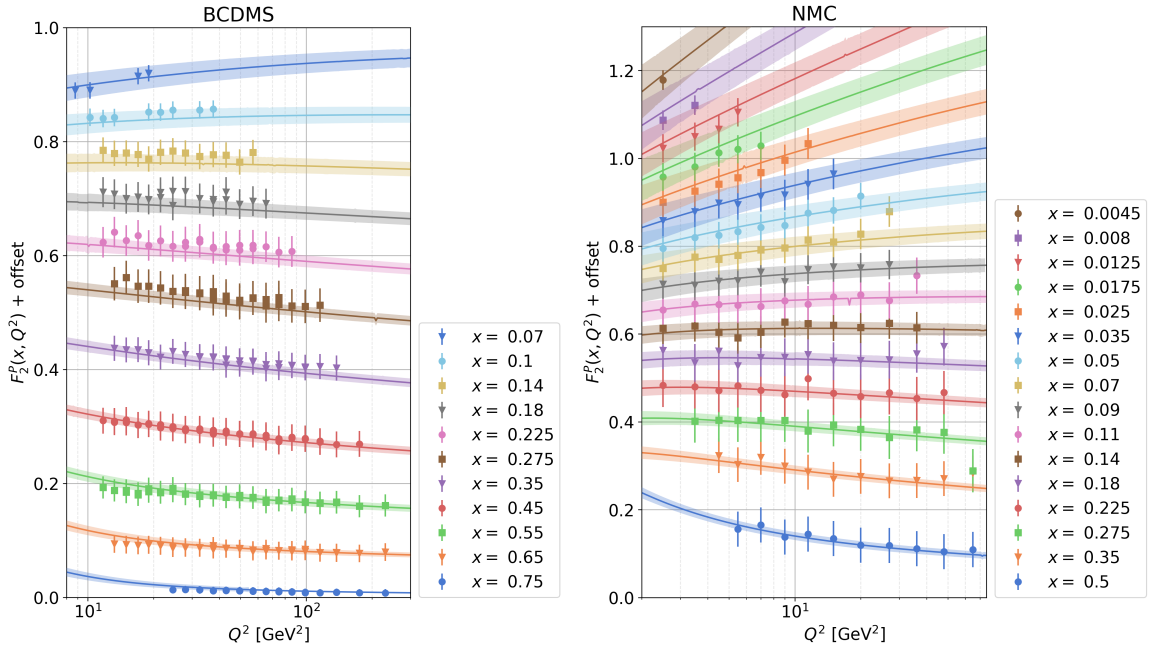


Figure 4.1: Kinematic coverage of experimental data.

We use the F123 module of the nCTEQ codebase to calculate  $F_2$  from the best-fit and each eigenvector PDF. Figure 4.2 shows the result for  $F_2$  as solid lines with shading indicating the area between  $F_2 + \Delta F_2$  and  $F_2 - \Delta F_2$ . The  $F_2$  data for each value of  $x$  are separated by an offset of  $0.05i$  where  $i$  counts through the  $x$  bins in descending order, starting at  $i = 0$ .

Figure 4.2: Comparison of data and theoretical prediction for  $F_2$  data from the BCDMS group (left) and the NMC group (right). The offset between  $x$  bins is  $0.05i$  where  $i$  counts through the bins in descending order, starting at 0.

The theory agrees very well with the data, which should come as no surprise since the data sets are part of the CTEQ6.6 fit. The uncertainties are fairly small in the high  $x$  region for both the data and theory, but in the very small  $x$  region the scarcity of data leads to large uncertainties. The violation of Bjorken scaling can be seen in both data sets in Figure 4.2, even though it is rather weak in large parts of this kinematic range. It is displayed most clearly in the data for  $0.0125 \leq x \leq 0.07$  but the theory values predict that data in the very small  $x$  region for  $Q^2$



values higher than the available ones would display the effects of the DGLAP evolution even stronger.

## 4.2 PDF fitting

We now have all the required theory, tools and data to fit parton distribution functions ourselves. The final missing part is the functional form of the PDFs, for which we decide to use the nCTEQ15 parameterization.

### 4.2.1 nCTEQ parameterization

The parameterization of the PDFs is specified at a fixed low-energy scale  $Q_0^2$ , which we set to  $1.69 \text{ GeV}^2$ . PDFs at higher  $Q^2$  are determined through NLO perturbative QCD evolution. The nCTEQ15 functional form is given by

$$xf(x, Q_0^2) = a_0 x^{a_1} (1-x)^{a_2} e^{a_3 x} (1 + e^{a_4 x})^{a_5}, \quad (215)$$

with independent parameters for  $u_v = u - \bar{u}$ ,  $d_v = d - \bar{d}$ ,  $g$ ,  $\bar{u} + \bar{d}$ ,  $s + \bar{s}$ . The top quark is neglected entirely, while charm and bottom are assumed to be  $c(x, Q_0) = b(x, Q_0) = 0$ . For the ratio  $\frac{\bar{d}}{\bar{u}}$  the form

$$\frac{\bar{d}(x, Q_0^2)}{\bar{u}(x, Q_0^2)} = b_0 x^{b_1} (1-x)^{b_2} + (1 + b_3 x)(1-x)^{b_4} \quad (216)$$

is used. The normalization parameters  $a_0^{(i)}$  and  $b_0$  are not part of the fit as that would allow for potentially unphysical results that violate sum rules such as (88) - (90). The normalizations are therefore determined directly from the sum rules whenever the parameters change. Due to the limited data included in the fit we will also constrain ourselves to fitting only the valence quark distributions  $xu_v(x, Q^2)$  and  $xd_v(x, Q^2)$ . This leaves us with  $d = 10$  parameters for our fits and all other parameters are kept fixed at the value of the nCTEQ15 fit. Figure 4.3 shows the effect of changing each parameter to help get a feeling for the meaning of the five fit parameters in (215). Figure A.2 at the end of the appendix contains equivalent plots with logarithmic  $x$  scale to allow a better look at the small  $x$  region. We see that the parameter  $a_1$  mostly affects how  $xf(x)$  tends towards 0 as  $x$  does the same, as it is part of the factor  $x^{a_0}$ . However, normalizing the distributions via the sum rules makes it so that varying each individual parameter changes the shape of the distribution across the entire  $x$  range. The second parameter  $a_2$  belonging to the factor  $(1-x)^{a_2}$  has the most pronounced effect on the very high  $x$  region of all the parameters. It changes the shape from a steep descend towards 0 as  $x \rightarrow 1$  for low values of  $a_2$  to a flat region for large values. Increasing the fourth parameter moves the peak to the right and changes the behaviour near  $x = 0$  to make the function more flat. Decreasing the parameter moves the peak to the left but does not affect the region near  $x = 0$  as much as the other direction. Increasing or decreasing the parameter beyond the given interval does not change the shape much further as indicated by the fact that the outermost curves lie very close

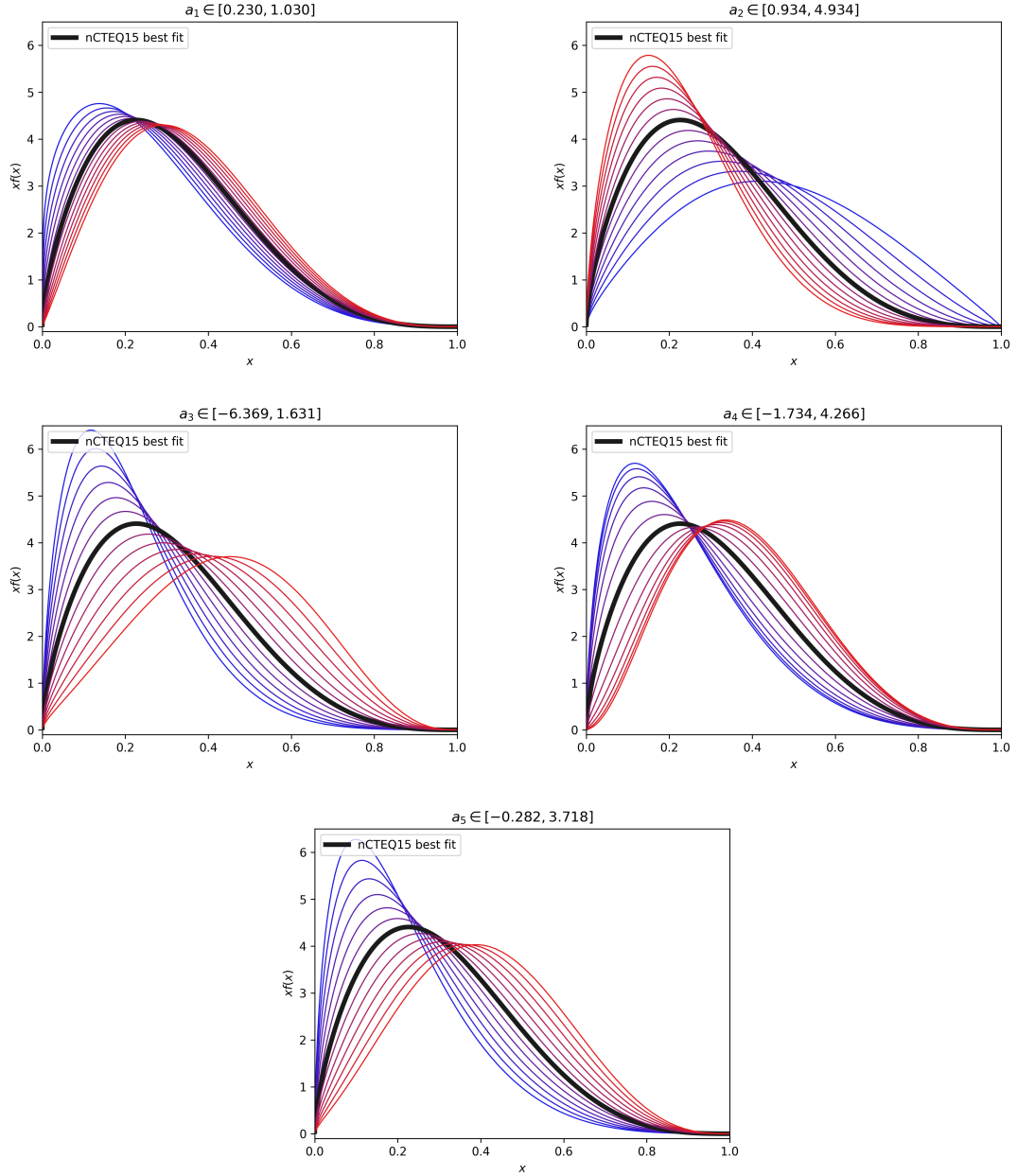


Figure 4.3: Comparison of the impact of each individual parameter on the shape of the parton distribution at the input scale  $Q_0^2 = 1.69 \text{ GeV}^2$ . The blue curves represent the lower end of the given interval and the red curve the upper end. Each curve represents an equidistant step in parameter space and the solid black curve shows the best fit curve from the nCTEQ15 fit [33]. Each curve is normalized to fulfil the sum rule (88).

together. The third and fifth parameter have very similar effects on the shape of the function if the other parameters are fixed at the values of the nCTEQ15 fit. They hardly change the behaviour for very low or very high  $x$  but instead move the position of the maximum towards the higher  $x$  values and make it wider as they are increased. The fact that they have such similar effects on the function might become problematic when fitting the PDFs to data because the direction in  $\chi^2$  space where one of the two is increased while the other is decreased accordingly will be very flat.

### 4.2.2 PDF fits

At this point all prerequisites for fitting PDFs are fulfilled. We start each fit with parameters that give a  $\chi^2$  value on the order of  $10^4$  ( $\frac{\chi^2}{\nu} \approx 10$ ) and then run the Levenberg-Marquardt algorithm using the  $K$  factor approximation. Then we use the full theory to repeat the algorithm using the parameters of the first run as initial values. As a final fitting step the improvement algorithm from section 3.3.5 is run. We repeat this for four different parameter combinations:

- 6 Parameters:  $\{a_1^{(u_v)}, a_2^{(u_v)}, a_5^{(u_v)}, a_1^{(d_v)}, a_2^{(d_v)}, a_5^{(d_v)}\}$
- 8 Parameters (a):  $\{a_1^{(u_v)}, a_2^{(u_v)}, a_3^{(u_v)}, a_5^{(u_v)}, a_1^{(d_v)}, a_2^{(d_v)}, a_3^{(d_v)}, a_5^{(d_v)}\}$
- 8 Parameters (b):  $\{a_1^{(u_v)}, a_2^{(u_v)}, a_4^{(u_v)}, a_5^{(u_v)}, a_1^{(d_v)}, a_2^{(d_v)}, a_4^{(d_v)}, a_5^{(d_v)}\}$
- 10 Parameters:  $\{a_1^{(u_v)}, a_2^{(u_v)}, a_3^{(u_v)}, a_4^{(u_v)}, a_5^{(u_v)}, a_1^{(d_v)}, a_2^{(d_v)}, a_3^{(d_v)}, a_4^{(d_v)}, a_5^{(d_v)}\}$

We choose  $a_1, a_2, a_5$  as the minimal parameterization because, as seen in the previous section, the first two parameters determine the low- and high- $x$  region. We chose  $a_5$  to parameterize the medium- $x$  region because  $a_4$  was shown in figure 4.3 to only affect the shape to a relatively small extent and there is hardly a difference in the effects of  $a_3$  or  $a_5$ . For the two 8 parameter sets we include either one of the remaining two parameters for both  $u_v$  and  $d_v$  and the 10 parameter set includes all parameters. The remaining parameters for  $u_v$ ,  $d_v$  and all other flavours are kept fixed at the values of the nCTEQ15 fit. The parameters of the resulting PDFs and their  $\chi^2$  values are given in table 2 along with the parameters and  $\chi^2$  of the nCTEQ15 fit.

Table 2: Parameters and  $\chi^2$  value of the four fits and the nCTEQ15 fit. The upper parameter in each cell is for  $u_v$ , the lower one for  $d_v$ . The cells where values remained fixed at the nCTEQ15 values are greyed out.

Fit	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$\chi^2$
nCTEQ15	0.6300 0.5133	2.9335 4.2114	-2.3692 -2.3745	1.2664 0.9650	1.7184 3.0000	1292.83
6 parameters	0.6112 0.5429	3.0384 4.5107	-2.3692 -2.3745	1.2664 0.9650	1.8666 2.8282	1188.43
8 parameters (a)	0.6282 0.7025	2.9536 10.3073	-2.3708 14.7610	1.2664 0.9650	0.7025 -3.4760	1185.44
8 parameters (b)	0.6212 0.5366	3.0297 5.2149	-2.3692 -2.3745	1.2300 0.5130	1.8743 4.6590	1187.27
10 parameters	0.6242 0.5331	3.0439 7.7468	-1.7325 4.4962	1.5296 0.2900	1.3055 2.1484	1184.23

For the most part the minima found by our algorithm lie in the close vicinity of the nCTEQ15 parameters. One notable exception are the  $d_v$  parameters in the first 8 parameter fit, where the parameters do not look like the other sets at all. We predicted that this might happen when we looked at the effects of changing each parameter in 4.3 and saw how similar  $a_3$  and  $a_5$  looked, though it seems that  $a_1$  and  $a_2$  also vary somewhat along this valley in  $\chi^2$  space.

The consistent value of  $\chi^2$  between the various fits tells us both that 6 parameters are sufficient to fit the given data and that we are unlikely to be far away from the "true" minimum. We

see our fit yielding somewhat better  $\chi^2$  values than the nCTEQ15 fit, which does not mean that we somehow found better PDFs than a vastly more complex global analysis, but rather that the inclusion of other data sets from different experiments moves the minimum to some extent. If our value of  $\chi^2$  would be significantly lower than the nCTEQ15 value it would indicate disagreement between the BCDMS / NMC experiments and the other experiments used in the global fit.

Plugging the  $\chi^2$  values into (208) we get roughly  $T = 7.32$  for each parameterization, which we round up slightly to  $T = 7.5$ . The corresponding  $t$  values for  $N_{\text{params}} = \{6, 8, 10\}$  are then calculated from (211) to be  $t = \{3.1, 2.7, 2.4\}$ . The final PDFs with the uncertainty bands calculated via the Hessian method are shown in figure 4.4 at the initial scale  $Q_0^2 = 1.69 \text{ GeV}^2$ . Figure 4.5 shows the uncertainties normalized by the values of the corresponding quark distributions.

The four fits have very similar central predictions even though we saw the parameters looked rather dissimilar. It would have been problematic if the different parameters had actually lead to differently shaped PDFs as that would mean that the data could not properly constrain them. The fact that the parameters can be changed without affecting the shape of the distribution too much is a problem in itself however. Looking at figure 4.5 we see the uncertainties rising when more parameters are added because the minima lying in flat valleys in  $\chi^2$  space allow large variations in parameters. The two 8 parameter fits show that including  $a_3$  causes a larger rise the region around  $x = 0.1$  where it has a very similar effect as  $a_5$ . The inclusion of  $a_4$  also leads to rising uncertainties but the medium  $x$  region is less affected than it was by  $a_3$  and instead the low  $x$  region becomes less certain. The fit with 10 parameters has the largest uncertainties across the entire  $x$  region because the limited data from just two DIS experiments cannot properly constrain this amount of parameters. For all fits the uncertainties are high in the low  $x$  region and diverge quickly as  $x \rightarrow 1$ , which reflects the kinematic region of the data where we have no information for  $x$  values above 0.75 and only a few data points below  $x = 0.01$ . Generally we see the lowest uncertainties around  $x = 0.1$ , where the bulk of the data is located. The relative uncertainties for the  $d_v$  are also somewhat lower than for  $u_v$ , which might stem from their lower electric charge and the associated weaker contribution to  $F_2$ .

Figures 4.6 and 4.7 show the PDFs from the 6 parameter fit at various scales and their relative uncertainties respectively. The behavior of the valence quark distributions under variation of the scale looks exactly like it did for the nCTEQ15 version in figure 2.11. This should come as no surprise considering the fact that we chose the same parameterization and our parameters are relatively close. The scale hardly affects the uncertainties in this case aside from some slightly enhanced uncertainties at the lowest scale, which lies outside the kinematic region of our data. We would see bigger changes in uncertainties if we had included more parameters of other partons in our fit, because the PDFs of the different flavours mix under DGLAP evolution. If we had one flavour with large uncertainties in the small  $x$  region for example we would expect this uncertainty to propagate into the small  $x$  region of the  $u_v$  and  $d_v$  distributions as the scale increases. Since the uncertainties of  $u_v$  and  $d_v$  are qualitatively very similar and we have excluded all other flavours and gluons from the fit we have no reason to expect qualitative changes in the uncertainties as we change the scale.

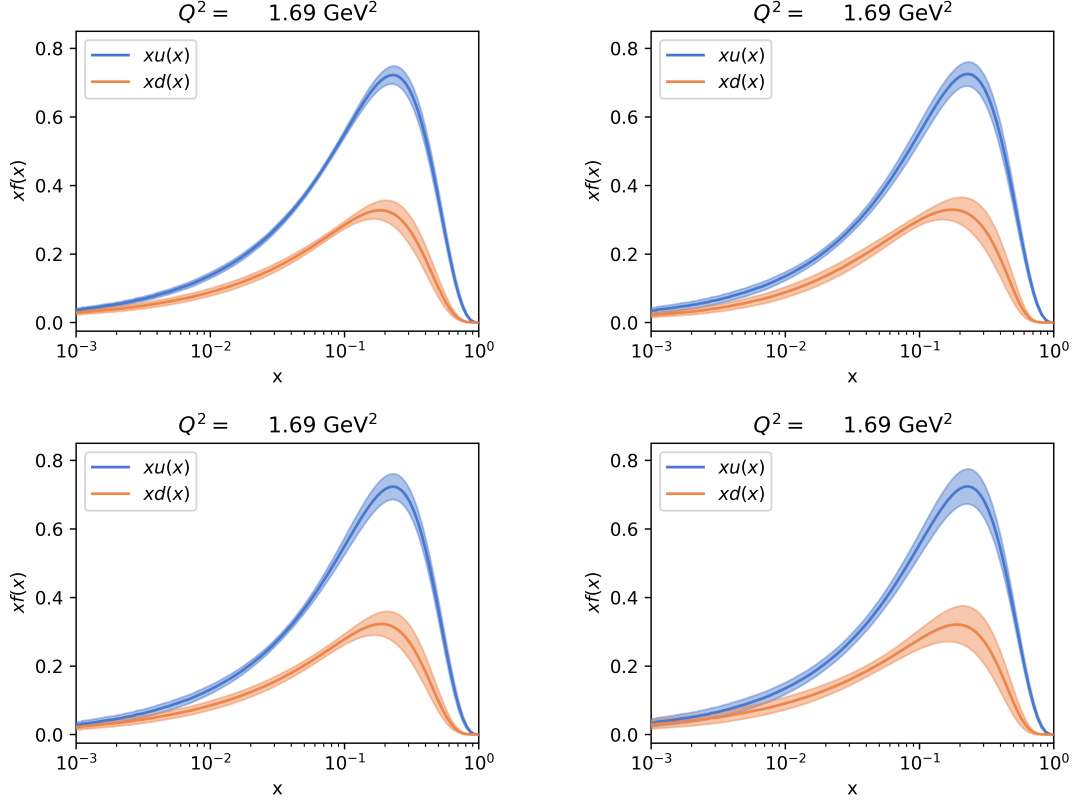


Figure 4.4: Comparison between the valence quark distributions of the four fits at input scale  $Q_0^2 = 1.69 \text{ GeV}^2$  (upper left: 6 parameters, upper right: 8 parameters (a), lower left: 8 parameters (b), lower right: 10 parameters). The eigenvector curves have been left out for visual clarity.

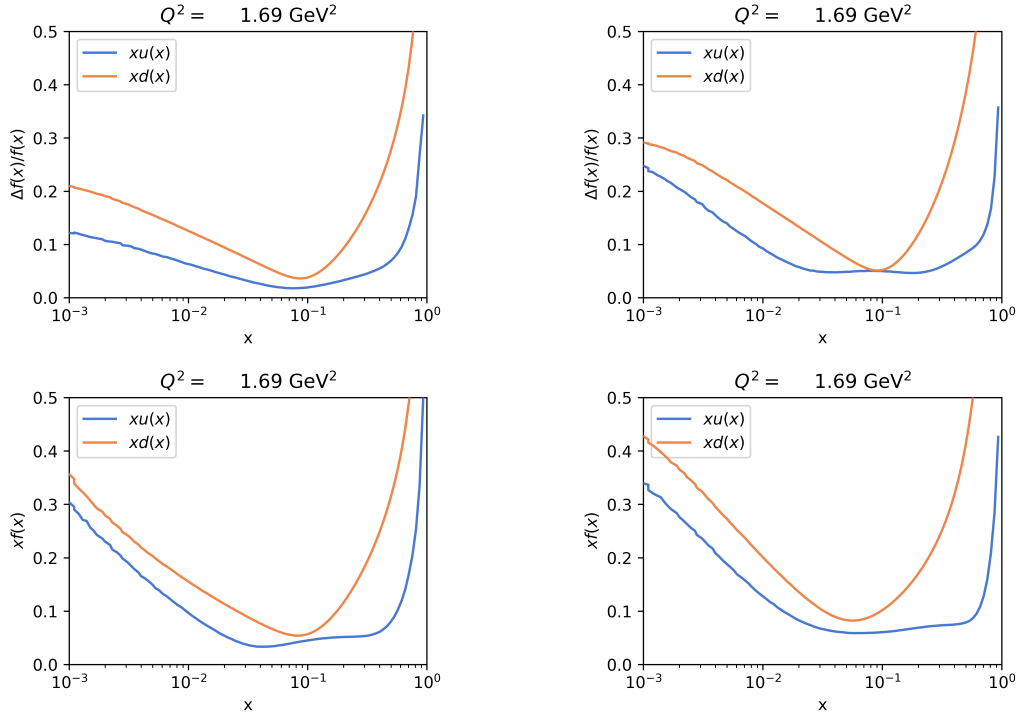


Figure 4.5: Comparison between the uncertainties of the four fits at input scale  $Q_0^2 = 1.69 \text{ GeV}^2$  (upper left: 6 parameters, upper right: 8 parameters (a), lower left: 8 parameters (b), lower right: 10 parameters).

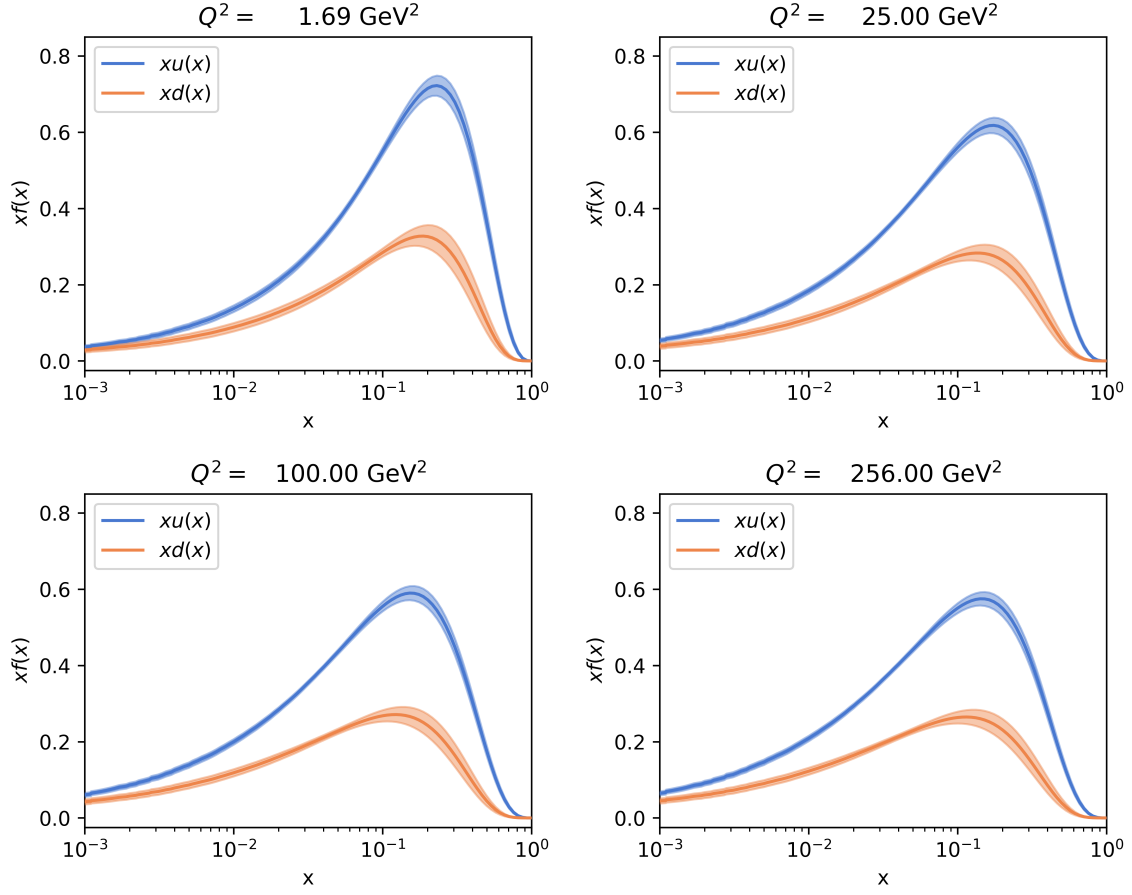


Figure 4.6: Comparison between the valence quark distributions of the six parameter fit at different scales.

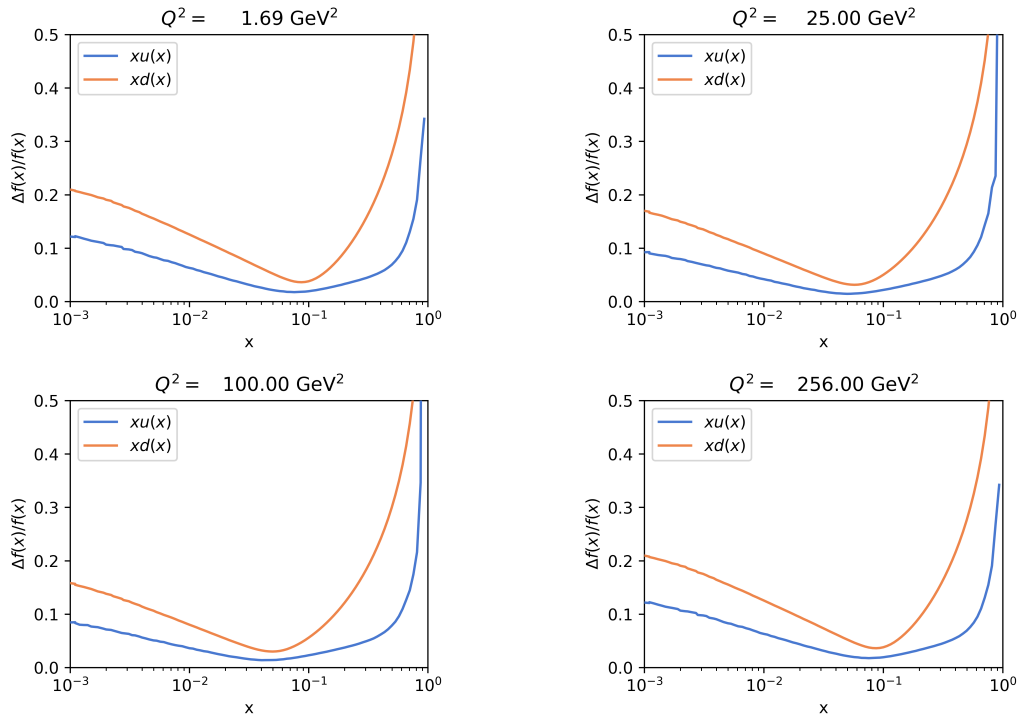


Figure 4.7: Comparison between the uncertainties of the 6 parameter fit at different scales.

We can also get a measure of both the fit quality and the Hessian method's quadratic approximation by plotting the  $\Delta\chi^2_i(z)$  curves for each eigenvector direction  $M_i$ . If the quadratic approximation was perfect and the fit exactly at the minimum we would expect all curves to be unit parabolas. Figure 4.8 shows that all curves are reasonably close to the unit parabola drawn in black. In the cases of the 6 and 8 parameter fit we can see that some curves are not perfectly centered. The reductions in  $\chi^2$  lie on the order of  $\frac{\Delta\chi^2}{\nu} < 10^{-4}$ , however, which means they are close to the limit of numerical precision caused by the numerical DGLAP evolution. Further improvements of the PDFs would therefore have to take this numerical precision into account for example by using the improved matrix  $M$  in the final fit improvement algorithm. Because  $M$  needs to be evaluated anew after each iteration of the algorithm however, this would increase the computation time by a huge amount for an improvement in  $\chi^2$  that is two orders of magnitude smaller than the uncertainty  $T^2 = 56.25$ . Small deviations from the minimum are also automatically accounted for in the uncertainties because we chose the normalization (214) for the eigenvectors. We know that slight improvements of the fit in this region will not open up the path to steeper regions in  $\chi^2$  space with even lower values, because we have the 10 parameter fit sitting nicely inside the center and it would not make sense to have a 6 or 8 parameter fit yield better  $\chi^2$  values.

The quadratic approximation also fits well with the shapes of the  $\Delta\chi^2$  curves even though some show slight contributions from higher order terms, most notably the pink and violet curves in the 10 parameter fit. Because we chose our normalization of the curves so that the lower of the two  $\Delta\chi^2(z = \pm T)$  is equal to  $T^2$  and we have no curves with shapes that are completely incompatible with the quadratic approximation we have made sure that no uncertainties in the final result are underestimated.

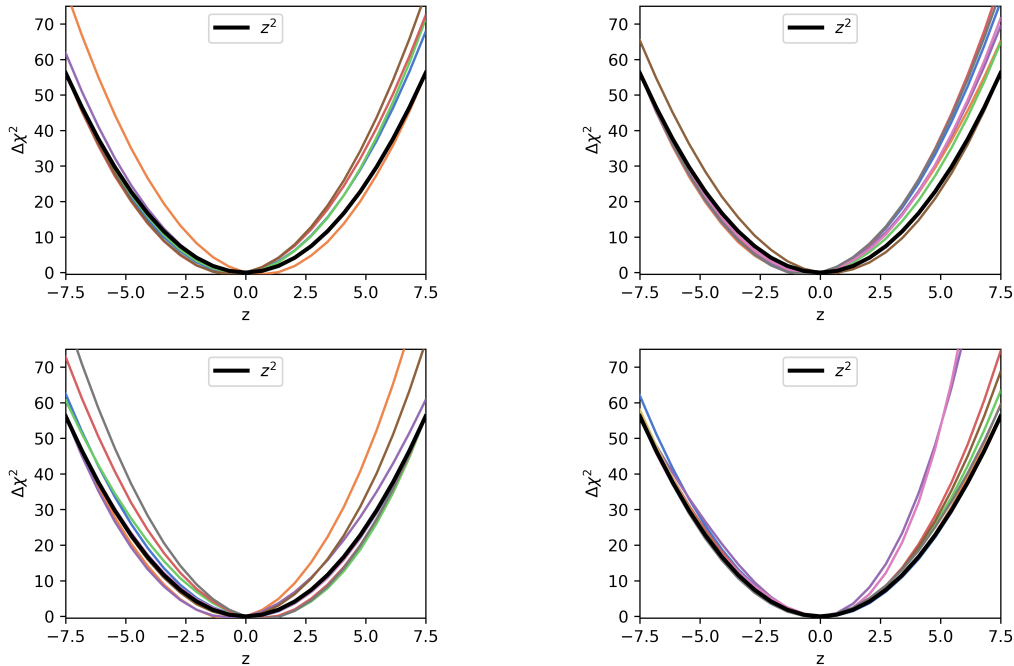


Figure 4.8: Comparison between the  $\Delta\chi^2$  curves along all eigenvectors of the four fits (upper left: 6 parameters, upper right: 8 parameters (a), lower left: 8 parameters (b), lower right: 10 parameters).

### 4.2.3 Comparison with data

As a final test we want to compare our fit with the data. To do this we calculate  $F_2$  using the module F123 of the nCTEQ code for each set of PDFs. With the corresponding sets of eigenvector PDFs we can compute the uncertainties using (202). Figure 4.9 shows the comparison of the four fits with the BCDMS data while figure 4.10 shows the same for the NMC data. The figures use the same offset between  $x$  bins as figure 4.2 and the region excluded by the kinematic cuts is marked in the plots as a hatched area (not visible in the BCDMS plot, because no data was cut there).

Overall we see very good agreement of the central values of the prediction with the data aside from some slight disagreements with a few NMC data points, in the high  $Q^2$  end of the  $x = \{0.07, 0.11, 0.275\}$  data. These seem to be statistical outliers however instead of physically meaningful deviations as the corresponding data of the BCDMS experiment does not show any such features. The data in the excluded region mostly agrees with the prediction aside from the small  $x$  region where the data seems to consistently lie slightly above the prediction. These disagreements are likely due to higher twist effects, that are especially relevant in this region and they were the reason why we excluded the low  $Q^2$  region in the first place.

The uncertainties of the  $F_2$  prediction show the same behavior with regards to variation of  $x$  and  $Q^2$  as seen in figures 4.4 and 4.5. The reason for that is simply that we did not include any other partons besides the valence quarks in our fit parameters. This also leads to an important caveat to be made with regards to the uncertainties in the small  $x$  region, which is dominated by the sea quark contributions. The gluon and sea quark parameters however remained fixed at the nCTEQ15 values and therefore they are ignored by the Hessian method. If we had included parameters of the  $\bar{u} + \bar{d}$  or  $g$  distribution in our fit we would likely see behavior similar to that of the CTEQ6.6 fit in figure 4.2, where the uncertainties rise as  $x$  gets lower and data less abundant. This is one of the weaknesses of the Hessian method because there is no straightforward way to include the uncertainties of parameters that are relevant for the result but remain fixed during the fit. As we have already mentioned in the previous section, more open parameters would require the inclusion of more data sets and particularly data sets from processes other than DIS to make a larger kinematic region available.



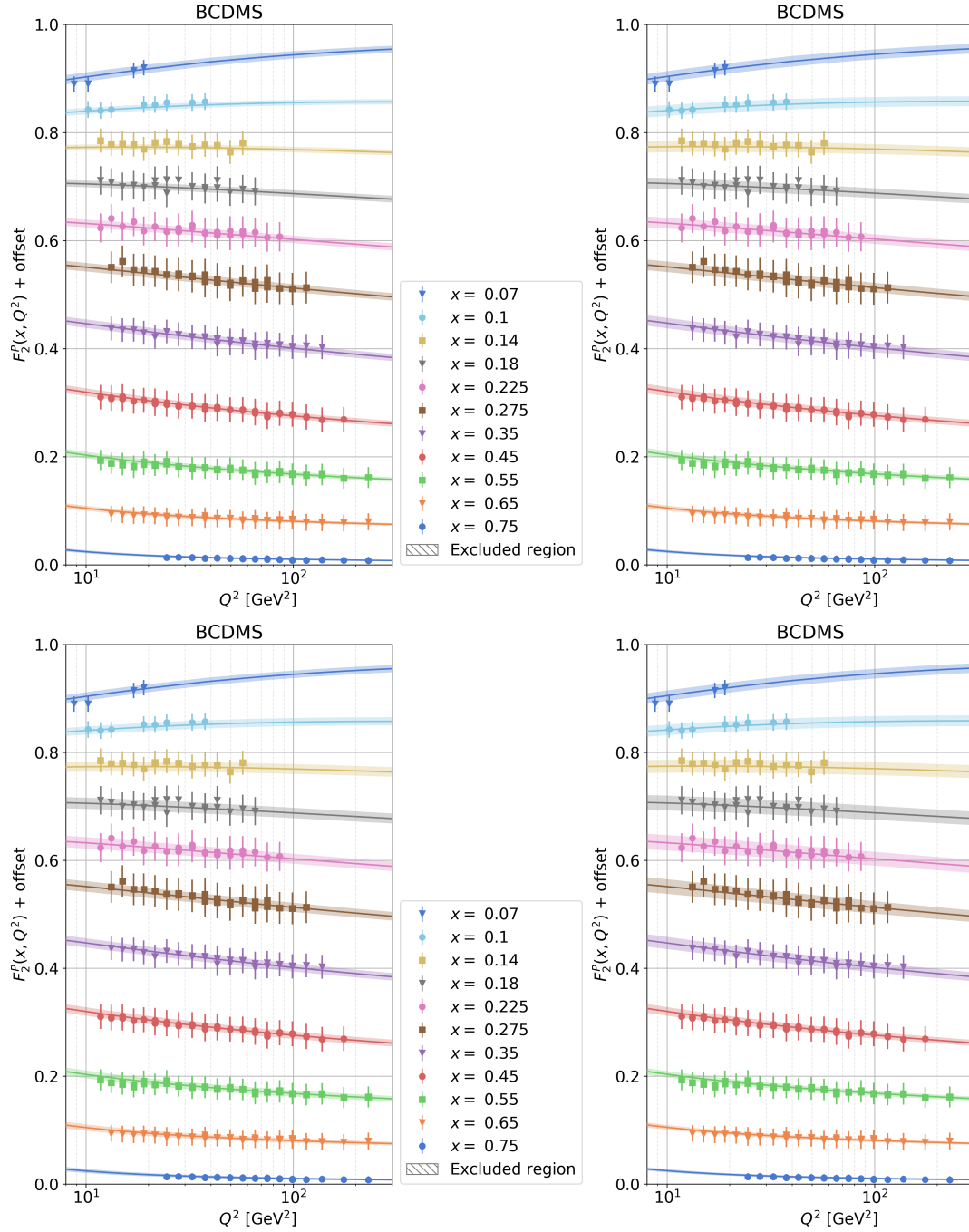


Figure 4.9: Comparison between  $F_2(x, Q)$  from the experimental data of the BCDMS group and the prediction from the four fits (upper left: 6 parameters, upper right: 8 parameters (a), lower left: 8 parameters (b), lower right: 10 parameters).

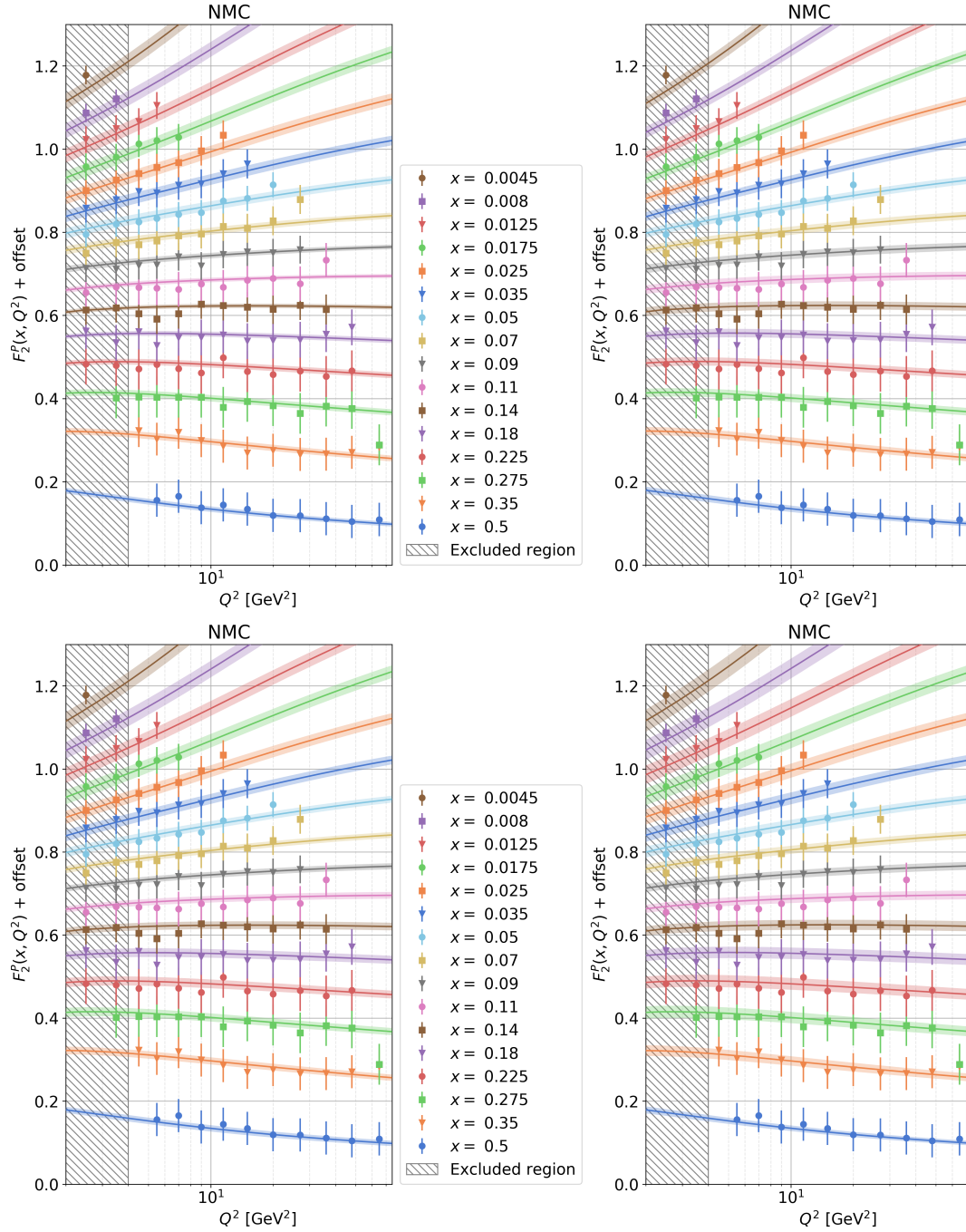


Figure 4.10: Comparison between  $F_2(x, Q)$  from the experimental data of the NMC group and the prediction from the four fits (upper left: 6 parameters, upper right: 8 parameters (a), lower left: 8 parameters (b), lower right: 10 parameters).

## 5 Summary and outlook

In this thesis all the necessary groundwork to understand parton distribution functions and their determination from experimental data was presented. In addition, further detail went into the some key aspects to reveal some of the intricacies that would remain hidden when talking only in terms of general concepts.

Firstly, the cross section for various scattering processes at leading order were derived and then the parton model was introduced from the experimental observation that at high energies the scattering cross section of an electron on a proton becomes that of an electron scattering on a point-like particle. As part of the parton model's discussion, the phenomenon of Bjorken scaling was introduced and the relation between the proton structure functions  $F_2$  and the constituent partons was explored. Building upon this the NLO-QCD corrections to the parton model were computed, which result in the DGLAP evolution equations, whose discussion concluded the theory part.

The next section introduced the Levenberg-Marquardt algorithm as the choice of fitting method and discussed some improvements and adaptations necessary for usage with PDFs. To complement the Levenberg-Marquardt algorithm which uses the Hessian matrix in each iteration, the Hessian method was chosen to estimate the uncertainties of our fits. After laying out the formalism of the Hessian method, an algorithm was introduced to improve the fit by building upon this formalism and everything was tried on some illustrative example fits.

The methods were then integrated into the nCTEQ codebase to make use of their implementation of the theory computations and to allow easy use and extensibility by others in the future. The last section covered the determination of PDFs for the valence quarks  $u_v$  and  $d_v$  using data from the BCDMS and NMC experiments. The fits were computed with different numbers of fit parameters and comparisons were done between the individual fits. Finally, some tests were performed on the validity of the fit and the assumptions of the Hessian method including the calculation of predictions for  $F_2$  from the fits to compare with the original data.

Because of the broad scope of ingredients that go into the determination of PDFs, the work done in this thesis can be expanded in a wide variety of ways. With the integration of the methods into the nCTEQ codebase the groundwork is already laid for the incorporating of more processes and the corresponding experimental data. This would also open up the possibility of including more parameters in the fit or trying entirely new parameterizations. The implementation of further fitting algorithms like Markov chain Monte Carlo methods or additional methods of uncertainty estimation like resampling via bootstrapping / jackknifing could be used to validate the results obtained from the methods presented in this work and to draw comparisons in runtime.

## A Appendix

The appendix contains short sections about topics relevant to the theory part of the thesis like summaries of regularization and QCD-relevant group theory aspects. It also includes many identities and relations that are only referenced in the main text to keep it focussed on the topics at hand without distraction. Interesting or complex relations like the plus distribution identity and the integral (A.23) are given with derivation.

### A.1 Summaries and derivations

#### A.1.1 Regularization

When performing calculations of loop diagrams needed for higher-order corrections one often runs into divergent integrals like

$$\int \frac{d^4 q}{(q^2 - \Delta + i\epsilon)^2}. \quad (\text{A.1})$$

To be able to handle these divergencies, we need to know their behaviour in an explicit form rather than just knowing that these integral diverge. Various methods of regularization have been developed deal with this problem.

One approach would be the introduction of a hard cutoff on the energy component of the loop momentum  $q_E < \Lambda$ . The result of the integral then depends on the choice of the cutoff scale  $\Lambda$  and the divergence is still present in the limit  $\Lambda \rightarrow \infty$ . The problem with this method is the breaking of Lorentz invariance as the cut is only applied to one component of the momentum vector.

There are various other methods of regularization like Pauli-Villars regularization, where for each physical particle a *ghost* particle of mass  $\Lambda$  is added with wrong statistics or a kinetic term with wrong sign so that the loop amplitudes of the ghosts and the physical particles cancel at asymptotically large loop momentum.

The method we employ in this thesis is called *dimensional regularization* and is best explained with an example:

$$\int_a^\infty \frac{1}{q^2} d^3 q \quad \text{---} \quad \text{linearly divergent}, \quad (\text{A.2})$$

$$\int_a^\infty \frac{1}{q^2} d^2 q \quad \text{---} \quad \text{logarithmically divergent}, \quad (\text{A.3})$$

$$\int_a^\infty \frac{1}{q^2} dq \quad \text{---} \quad \text{convergent}. \quad (\text{A.4})$$

The behaviour of the divergence changes with the dimension of the integration. Lowering the dimension of an ultraviolet (UV) divergent integral can make it finite. Therefore the principal idea of dimensional regularization is evaluating all loop integrals in  $D = 4 - 2\epsilon$  space-time dimensions, where  $\epsilon > 0$  for UV divergencies and by choosing  $\epsilon < 0$  we can also handle in-

frared (IR) divergencies. The result of the integrals then depends on the dimension  $D$  and the divergence emerges when  $\epsilon \rightarrow 0$ . When computing a normally divergent integral with a cutoff scale one can see the degree of divergence in the powers of the cutoff scale in the final result. In dimensional regularization however, the degree of divergence of the original integral is encoded in the structure of the singularities in the  $D$ -dimensional result. For example the integral  $\int \frac{d^D q}{(q^2 - \Delta)^2} \propto \Gamma\left(\frac{4-D}{2}\right)$  has a pole at  $D = 4$  and is logarithmically divergent in  $D = 4$  dimensions. The integral  $\int \frac{d^D q}{q^2 - \Delta} \propto \Gamma\left(\frac{2-D}{2}\right)$  has poles at  $D = 4$  and  $D = 2$  and is quadratically divergent in  $D = 4$  dimensions.

Dimensional regularization leaves Lorentz invariance intact, but requires us to formulate the entire quantum field theory in  $D$  dimensions. Most importantly this means that four-vectors become  $D$ -vectors

$$q^\mu = (q^0, q^1, q^2, q^3) \rightarrow (q^0, \dots, q^{D-1}) \quad (\text{A.5})$$

and that we have  $D$  Dirac matrices instead of four. How this affects the various properties of the Dirac matrices can be found in section A.2.1, while sections A.1.2 and A.1.3 explain how to solve some common integrals in  $D$  dimensions.

The Lagrangian density in  $D$  dimensions needs mass dimension  $D$  so that the action is dimensionless. For the QED Lagrangian

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)^2 + \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - e\bar{\psi}\gamma^\mu \psi A_\mu, \quad (\text{A.6})$$

this implies

$$[A_\nu] = \frac{D-2}{2}, \quad [\psi] = \frac{D-1}{2}, \quad [m] = 1, \quad [e] = \frac{4-D}{2}. \quad (\text{A.7})$$

It is convention to take

$$e \rightarrow \mu^{\frac{4-D}{2}} e, \quad (\text{A.8})$$

with  $[\mu] = 1$  to retain the original dimensionless coupling.

### A.1.2 $D$ -dimensional spherical coordinates

When we encounter a spherically symmetric integral in  $D$ -dimensions, we can use

$$\int d^D k = \int d\Omega_D \int k^{D-1} dk, \quad (\text{A.9})$$

where  $d\Omega_D$  denotes the differential solid angle of a  $D$ -dimensional unit-sphere:

$$d\Omega_D = \sin^{D-2}(\phi_{D-1}) \sin^{D-3}(\phi_{D-2}) \dots \sin(\phi_2) d\phi_1 \dots d\phi_{D-1}. \quad (\text{A.10})$$

The interval of integration for the angle to the  $i$ -th axis  $\phi_i$  is  $[0, \pi]$  for  $i > 1$  and  $[0, 2\pi]$  otherwise. The surface area of a  $D$ -dimensional unit sphere is  $\Omega_D = \int d\Omega_D$  and can be calculated in the

following way:

$$(\sqrt{\pi})^D = \left( \int_{-\infty}^{\infty} dx e^{-x^2} \right)^D = \int d\Omega_D \int_0^{\infty} dr r^{D-1} e^{-r^2}. \quad (\text{A.11})$$

with the definition of the  $\Gamma$  function (A.104):

$$(\sqrt{\pi})^D = \frac{1}{2} \Gamma\left(\frac{D}{2}\right) \int d\Omega_D \quad (\text{A.12})$$

$$\Rightarrow \quad \Omega_D = \frac{2\pi^{\frac{D}{2}}}{\Gamma(\frac{D}{2})}. \quad (\text{A.13})$$

### A.1.3 Wick rotation and generic integrals

In loop calculations one often encounters integrals of the form

$$I_n(A) = \int \frac{d^D q}{(2\pi)^D} \frac{1}{(q^2 - A)^n}. \quad (\text{A.14})$$

The integration can be split into the time component  $q_0$  and the remaining  $D - 1$  space components. The integrand has poles in the complex plane:

$$q^2 - A + i\epsilon = 0 \quad \Rightarrow \quad q_0^2 - \vec{q}^2 - A + i\epsilon = 0 \quad \Rightarrow \quad q_0 = \pm \sqrt{\vec{q}^2 + A - i\epsilon}. \quad (\text{A.15})$$

Using the Cauchy integral theorem with a contour as shown in A.1 shows that the integral along

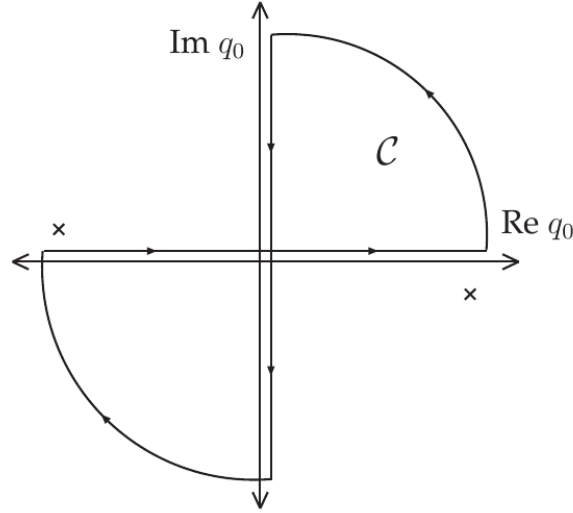


Figure A.1: Integral along the contour  $\mathcal{C}$  has to be 0, because both poles (marked with small crosses) are outside of the contour. Figure taken from [42].

the real axis is equal to the integral along the imaginary axis in negative direction, because both poles lie outside of the contour. This means we can write our integral as

$$I_n(A) = \int_{-\infty}^{\infty} dq_0 \int \frac{d^{D-1} q}{(2\pi)^D} \frac{1}{(q^2 - A + i\epsilon)^n} = \int_{-i\infty}^{i\infty} dq_0 \int \frac{d^{D-1} q}{(2\pi)^D} \frac{1}{(q^2 - A + i\epsilon)^n}. \quad (\text{A.16})$$

Now we introduce the momentum  $q_E$  with Euclidean metric  $q_E^2 = q_{E,0}^2 + \vec{q}_E^2$  as

$$q_0 = iq_{E,0}, \quad \vec{q} = \vec{q}_E \quad \longrightarrow \quad q^2 = q_0^2 - \vec{q}^2 = -q_{E,0}^2 - \vec{q}_E^2 = -q_E^2 \leq 0. \quad (\text{A.17})$$

Expressed in  $q_E$  the integral then becomes spherically symmetric:

$$I_n(A) = i \int_{-\infty}^{\infty} dq_{E,0} \int \frac{d^{D-1}q_E}{(2\pi)^D} \frac{1}{(-q_E^2 - A + i\epsilon)^n} = i \int \frac{d^D q_E}{(2\pi)^D} (-1)^n \frac{1}{(q_E^2 + A - i\epsilon)^n}. \quad (\text{A.18})$$

To solve this we introduce spherical coordinates in  $D$  dimensions as explained in the previous section:

$$\int d^D q_E = \int d\Omega_D \int_0^{\infty} q_E dq_E^{D-1} = \frac{1}{2} \int d\Omega_D \int_0^{\infty} dq_E^2 (q_E^2)^{\frac{D}{2}-1}. \quad (\text{A.19})$$

Plugging this into (A.18) gives

$$I_n(A) = i(-1)^n \frac{\pi^{\frac{D}{2}}}{\Gamma(\frac{D}{2})} \int_0^{\infty} \frac{dq_E^2}{(2\pi)^D} \frac{(q_E^2)^{\frac{D}{2}-1}}{(q_E^2 + A - i\epsilon)^n}. \quad (\text{A.20})$$

Now we substitute

$$y = \frac{A - i\epsilon}{q_E^2 + A - i\epsilon} \quad \Longrightarrow \quad dy = -\frac{A - i\epsilon}{(q_E^2 + A - i\epsilon)^2} dq_E^2, \quad (\text{A.21})$$

so that our integral becomes

$$\begin{aligned} I_n(A) &= i(-1)^n \frac{\pi^{\frac{D}{2}}}{\Gamma(\frac{D}{2})} \frac{1}{(2\pi)^D} \int_1^0 dy \frac{(q_E^2 + A - i\epsilon)^2}{A - i\epsilon} (-1) \left( \frac{A - i\epsilon}{y} - A + i\epsilon \right)^{\frac{D}{2}-1} \left( \frac{A - i\epsilon}{y} \right)^{-n} \\ &= i(-1)^n \frac{\pi^{\frac{D}{2}}}{\Gamma(\frac{D}{2})} (A - i\epsilon)^{\frac{D}{2}-n} \frac{1}{(2\pi)^D} \int_0^1 dy y^{n-1-\frac{D}{2}} (1-y)^{\frac{D}{2}-1}. \end{aligned} \quad (\text{A.22})$$

Using the definition of the  $B$ -function (A.105) and its relation to the  $\Gamma$ -function (A.106) we get our result:

$$I_n(A) = (-1)^n \frac{i}{(4\pi)^{\frac{D}{2}}} (A - i\epsilon)^{\frac{D}{2}-n} \frac{\Gamma(n - \frac{D}{2})}{\Gamma(n)}. \quad (\text{A.23})$$

When using the definition of the  $B$ -function, technically we have assumed that  $D > 2n$ . For  $D \rightarrow 4$  this means divergent results for  $I_1$  and  $I_2$ . These divergencies are responsible for the UV-divergences in many calculations in dimensional regularization.

In a similar fashion, one can derive

$$\int \frac{d^D q}{(2\pi)^D} \frac{q^2}{(q^2 - A)^n} = \frac{i}{(4\pi)^{D/2}} (-1)^{n-1} (D/2) \frac{\Gamma(n - D/2 - 1)}{\Gamma(n)} \left( \frac{1}{A} \right)^{n-D/2-1}. \quad (\text{A.24})$$

This also gives us the solution to another common integral in loop calculations:

$$\int d^D q \frac{q^\mu q^\nu}{(q^2 - A)^n} = \frac{1}{D} g^{\mu\nu} \int d^D q \frac{q^2}{(q^2 - A)^n}, \quad (\text{A.25})$$

which has to hold because the left hand side is a Lorentz tensor and only depends on the scalar  $A$ . Therefore dimensional analysis tells us that we can replace  $q^\mu q^\nu$  by  $cq^2 g^{\mu\nu}$  for some number  $c$ . Contracting both with  $g^{\mu\nu}$  tells us that  $c = \frac{1}{D}$ .

#### A.1.4 Mellin transform and Mellin convolution

The Mellin transform of a function is defined as

$$\{\mathcal{M}f(x)\}(n) = \int_0^\infty dx x^{n-1} f(x) = \tilde{f}(n). \quad (\text{A.26})$$

The inverse transform is

$$\{\mathcal{M}^{-1}\tilde{f}(n)\}(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} dn x^{-n} \tilde{f}(n) = f(x), \quad (\text{A.27})$$

where the integral is meant as a line integral in the complex plane that runs parallel to the imaginary axis and crosses the real axis in the value  $c$ .

The Mellin convolution is defined as

$$\{f *_M g\}(x) = \int_0^\infty \frac{d\xi}{\xi} f\left(\frac{x}{\xi}\right) g(\xi). \quad (\text{A.28})$$

In Mellin space this convolution reduces to a product:

$$\begin{aligned} \{\mathcal{M}(f *_M g)\}(n) &= \int_0^\infty dx x^{n-1} \int_0^\infty \frac{d\xi}{\xi} f\left(\frac{x}{\xi}\right) g(\xi) \\ &= \int_0^\infty dx x^{n-1} \int_0^\infty d\xi \int_0^\infty dz \delta(x - \xi z) f\left(\frac{x}{\xi}\right) g(\xi) \\ &= \int_0^\infty d\xi \int_0^\infty dz (\xi z)^{n-1} f(z) g(\xi) \\ &= \int_0^\infty d\xi \xi^{n-1} g(\xi) \int_0^\infty dz z^{n-1} f(z) \\ &= \tilde{f}(n) \tilde{g}(n). \end{aligned} \quad (\text{A.29})$$

#### A.1.5 Plus distribution

The plus distribution is defined through its integral multiplied by a test function:

$$\int_0^1 dz f(z) [g(z)]_+ = \int_0^1 dz (f(z) - f(1)) g(z) \quad (\text{A.30})$$

and

$$[g(z)]_+ = g(z) \quad \text{for } z \neq 1. \quad (\text{A.31})$$



Using this we can derive an important identity. We start with

$$\int_0^1 dz \frac{f(z)}{(1-z)^{1+\epsilon}} = \int_0^1 dz \frac{f(1)}{(1-z)^{1+\epsilon}} + \int_0^1 dz \frac{f(z) - f(1)}{(1-z)^{1+\epsilon}}, \quad (\text{A.32})$$

then use the definition of the beta function (A.105) on the first integral of the r.h.s. to get

$$\int_0^1 dz \frac{f(z)}{(1-z)^{1+\epsilon}} = -\frac{f(1)}{\epsilon} + \int_0^1 dz \frac{f(z) - f(1)}{(1-z)^{1+\epsilon}}. \quad (\text{A.33})$$

Expanding  $(1-z)^{-1-2\epsilon}$  as a series gives

$$\frac{1}{(1-z)^{1+\epsilon}} = \frac{1}{1-z} e^{-\epsilon \ln(1-z)} = \frac{1}{1-z} \sum_{n=0}^{\infty} \frac{(-\epsilon)^n}{n!} \ln^n(1-z). \quad (\text{A.34})$$

This can then be inserted into (A.33):

$$\int_0^1 dz \frac{f(z)}{(1-z)^{1+\epsilon}} = -\int_0^1 \frac{f(z)}{\epsilon} \delta(1-z) + \int_0^1 dz \sum_{n=0}^{\infty} \frac{(-\epsilon)^n}{n!} \ln^n(1-z) \frac{f(z) - f(1)}{(1-z)}. \quad (\text{A.35})$$

Plugging in the definition of the plus distribution (A.30) results in

$$\int_0^1 dz \frac{f(z)}{(1-z)^{1+\epsilon}} = -\int_0^1 \frac{f(z)}{\epsilon} \delta(1-z) + \int_0^1 dz \sum_{n=0}^{\infty} \frac{(-\epsilon)^n}{n!} f(z) \left[ \frac{\ln^n(1-z)}{(1-z)} \right]_+. \quad (\text{A.36})$$

Without the integral and the test function  $f(z)$  we get the desired identity

$$\frac{1}{(1-z)^{1+\epsilon}} = -\frac{1}{\epsilon} \delta(1-z) + \sum_{n=0}^{\infty} \frac{(-\epsilon)^n}{n!} \left[ \frac{\ln^n(1-z)}{(1-z)} \right]_+. \quad (\text{A.37})$$

### A.1.6 Feynman rules

#### Quantum electrodynamics

From the QED Lagrangian

$$\mathcal{L} = \frac{1}{4} F_{\mu\nu}^2 + i\bar{\psi} \not{D} \psi - m\bar{\psi} \psi \quad \text{with} \quad D_\mu \psi = \partial_\mu \psi + ie A_\mu \psi \quad (\text{A.38})$$

we get the following Feynman rules:

Photon propagator:

$$\text{~~~~~} = \frac{-i}{p^2 + i\epsilon} \left[ g_{\mu\nu} - (1-\xi) \frac{p_\mu p_\nu}{p^2} \right] \stackrel{\text{Feynman gauge}}{=} \frac{-ig_{\mu\nu}}{p^2 + i\epsilon}. \quad (\text{A.39})$$

Unless otherwise specified, we will always work in Feynman gauge.

Fermion propagator:

$$\longrightarrow = \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} \quad (\text{Arrow right for particles, left for antiparticles}). \quad (\text{A.40})$$

External photons (small circle denotes connection to the rest of the diagram):

$$\text{~~~~~}\bigcirc = \epsilon_\mu(p) \quad (\text{incoming}), \quad (\text{A.41})$$

$$\bigcirc\text{~~~~~} = \epsilon_\mu^*(p) \quad (\text{outgoing}). \quad (\text{A.42})$$

External fermions:

$$\longrightarrow\bigcirc = u^s(p), \quad (\text{A.43})$$

$$\bigcirc\longrightarrow = \bar{u}^s(p), \quad (\text{A.44})$$

$$\longleftarrow\bigcirc = \bar{v}^s(p), \quad (\text{A.45})$$

$$\bigcirc\longrightarrow = v^s(p). \quad (\text{A.46})$$

Interaction:

$$\begin{array}{c} \longrightarrow \\ \bullet \\ \text{~~~~~} \end{array} = \begin{array}{c} \longrightarrow \\ \bullet \\ \text{~~~~~} \end{array} = \begin{array}{c} \text{~~~~~} \\ \bullet \\ \begin{array}{c} \nearrow \\ \searrow \end{array} \end{array} = \begin{array}{c} \begin{array}{c} \nearrow \\ \searrow \end{array} \\ \bullet \\ \text{~~~~~} \end{array} = -ie\gamma^\mu. \quad (\text{A.47})$$

External fermions need to be on-shell according to the LSZ-reduction formula. This means we can use the equations of motion to simplify calculations:

$$(\not{p} - m)u^s(p) = \bar{u}^s(p)(\not{p} - m) = 0, \quad (\text{A.48})$$

$$(\not{p} + m)v^s(p) = \bar{v}^s(p)(\not{p} + m) = 0. \quad (\text{A.49})$$

## Quantum chromodynamics

The  $SU(N)$ -invariant Lagrangian of QCD for a set of  $N$  fermions and  $N$  scalars interacting with non-Abelian gauge fields is given by:

$$\begin{aligned} \mathcal{L} = & -\frac{1}{4}(F_{\mu\nu}^a)^2 - \frac{1}{2\xi}(\partial_\mu A_\mu^a)^2 + (\partial_\mu \bar{c}^a)(\delta^{ac}\partial_\mu + gf^{abc}A_\mu^b)c^c \\ & + \bar{\psi}_i(\delta_{ij}i\not{\partial} + gA^a_{ij}T_{ij}^a - m\delta_{ij})\psi_j, \end{aligned} \quad (\text{A.50})$$

$$\text{with} \quad F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc}A_\mu^b A_\nu^c. \quad (\text{A.51})$$

In the Lagrangian  $c^a$  and  $\bar{c}^a$  are the Fadeev-Popov ghosts and anti-ghosts respectively.

Gluon propagator:

$$\nu; b \text{ (wavy line)} \xrightarrow{p} \mu; a = \frac{-i}{p^2 + i\epsilon} \left[ g_{\mu\nu} - (1 - \xi) \frac{p_\mu p_\nu}{p^2} \right] \delta^{ab}. \quad (\text{A.52})$$

Ghost propagator:

$$b \text{ (dashed line)} \xrightarrow{p} a = \frac{-i}{p^2 + i\epsilon} \delta^{ab}. \quad (\text{A.53})$$

Colored fermion propagator:

$$b \text{ (solid line)} \xrightarrow{p} a = \frac{-i}{\not{p} - m + i\epsilon} \delta^{ab}. \quad (\text{A.54})$$

If a colored propagator appears as an intermediate state, we need to sum over all possible colors.

Fermion vertex:

$$\begin{array}{c} \mu; c \\ \text{(wavy line)} \\ \swarrow \quad \searrow \\ b \quad \quad a \end{array} = i\gamma^\mu T_{ab}^c. \quad (\text{A.55})$$

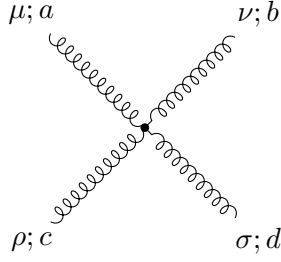
Ghost vertex:

$$\begin{array}{c} \mu; b \\ \text{(wavy line)} \\ \swarrow \quad \searrow \\ c^c \quad \quad \bar{c}^a \end{array} \quad p = -gf^{abc} p^\mu. \quad (\text{A.56})$$

3 gluon vertex:

$$\begin{array}{c} \nu; b \\ \text{(wavy line)} \\ \swarrow \quad \searrow \\ \rho; c \quad \quad \mu; a \end{array} \quad \begin{array}{c} p \\ \downarrow \\ k \end{array} = gf^{abc} [g^{\mu\nu} (k - p)^\rho + g^{\nu\rho} (p - q)^\mu + g^{\rho\mu} (q - k)^\nu]. \quad (\text{A.57})$$

4 gluon vertex:



$$\begin{aligned}
 &= -ig^2 \times [f^{abe} f^{cde} (g^{\mu\rho} g^{\nu\sigma} - g^{\mu\sigma} g^{\nu\rho}) \\
 &\quad + f^{ace} f^{bde} (g^{\mu\nu} g^{\rho\sigma} - g^{\mu\sigma} g^{\nu\rho}) \\
 &\quad + f^{ade} f^{bce} (g^{\mu\nu} g^{\rho\sigma} - g^{\mu\rho} g^{\nu\sigma})].
 \end{aligned} \tag{A.58}$$

### A.1.7 Group theory

#### Lie groups

A Lie group is a Group  $G$  where every element  $g \in G$  depends in a continuous and differentiable way on a set of real parameters  $\theta^a$ . Any element continuously connected to the identity can be written as

$$U = \exp(i\theta^a T^a) \cdot \mathbb{1}, \tag{A.59}$$

where  $T^a$  are the generators of the group. These generators form a Lie algebra

$$[T^a, T^b] = if^{abc} T^c. \tag{A.60}$$

The factors  $f^{abc}$  are called structure constants and  $[\cdot, \cdot]$  represents the Lie bracket which fulfills the following properties for any set of generators  $\{x, y, z\}$ :

- Bilinearity:  $[ax + by, z] = a[x, z] + b[y, z]$   
 $[z, ax + by] = a[z, x] + b[z, y]$
- Alternativity:  $[x, x] = 0$
- Anticommutativity:  $[x, y] = -[y, x]$
- Jacobi identity:  $[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0$

For our purposes the Lie bracket is just the commutator, as is the case in any physics application of Lie groups where the generators can be represented as matrices.

#### Representations and the group $SU(N)$

The special unitary group  $SU(N)$  is the Lie group of unitary  $N \times N$  matrices with determinant 1 and its group operation is the matrix multiplication. The  $N^2 - 1$  generators of the group form the corresponding Lie algebra  $\mathfrak{su}(N)$ .

A representation of a Lie group is an embedding of the group elements into operators acting on a vector space. Any finite dimensional representation is therefore comprised of matrices. A representation is called faithful if each group element gets its own matrix. The fundamental (or defining) representation is defined as the smallest faithful representation. For the algebra  $\mathfrak{su}(N)$  this means traceless hermitian  $N \times N$  matrices. Under infinitesimal group transformations they

act on a set of  $N$  fields  $\phi_i$  as

$$\phi_i \longrightarrow \phi_i + i\alpha^a (T_{\text{fund}}^a)_{ij} \phi_j \quad \text{with} \quad \alpha^a \in \mathbb{R}. \quad (\text{A.61})$$

The anti-fundamental representation acts on the complex conjugate fields  $\phi_i^*$  like the fundamental one acts on the fields  $\phi$ :

$$\phi_i^* \longrightarrow \phi_i^* + i\alpha^a (T_{\text{anti-fund}}^a)_{ij} \phi_j^* = \phi_i^* - i\alpha^a \phi_j^* (T_{\text{fund}}^a)_{ji}. \quad (\text{A.62})$$

Therefore the anti-fundamental generators can be expressed as  $T_{\text{anti-fund}}^a = -(T_{\text{fund}}^a)^*$ . Another important representation is the adjoint representation, which acts on the vector space spanned by the generators. Therefore the dimension of this representations has to be  $N^2 - 1$  for  $\text{SU}(N)$ . The  $(N^2 - 1) \times (N^2 - 1)$  matrices of the adjoint representation are given by

$$(T_{\text{adj}}^a)^{bc} = -if^{abc}. \quad (\text{A.63})$$

From now on we will use  $T^a \equiv T_{\text{fund}}^a$  for the generators in the fundamental representation and  $T_R^a$  for general representations.

#### Normalization:

The structure constants can be normalized arbitrarily, however due to (A.60) any chosen convention also determines the normalization of the generators. We choose to normalize the structure constants by

$$\sum_{c,d} f^{acd} f^{bcd} = N\delta^{ab}, \quad (\text{A.64})$$

which gives us

$$\text{Tr}(T^a T^b) = \frac{1}{2} \delta^{ab} \quad (\text{A.65})$$

for the generators. For  $\text{SU}(3)$  the generators in the fundamental representation are often written in a standard basis as

$$T_{\text{fund}}^a = \frac{1}{2} \lambda^a, \quad (\text{A.66})$$

where  $\lambda^a$  are the Gell-Mann matrices:

$$\begin{aligned} \lambda^1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda^2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda^3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda^4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \\ \lambda^5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, & \lambda^6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, & \lambda^7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, & \lambda^8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \end{aligned}$$

#### Basis independent properties:

Casimir operators or Casimirs are operators that characterize a representation independent of a basis. The most important one is the quadratic Casimir  $C_2(R)$ , which is defined by

$$T_R^a T_R^a = C_2(R) \mathbb{1}. \quad (\text{A.67})$$

The Dynkin index  $T(R)$  of a representation can be calculated from

$$\text{Tr}(T_R^a T_R^b) = T(R) \delta^{ab}. \quad (\text{A.68})$$

From our normalizations (A.64) and (A.65) we can easily see

$$T_F \equiv T(\text{fund}) = \frac{1}{2} \quad \text{and} \quad T_A \equiv T(\text{adj}) = N. \quad (\text{A.69})$$

Using (A.68) with  $a = b$ , we can evaluate the quadratic Casimir:

$$d(R) C_2(R) = T(R) d(G), \quad (\text{A.70})$$

where  $d(R)$  and  $d(G)$  are the dimension of the representation and the group respectively ( $d(\text{fund}) = N$  and  $d(\text{adj}) = d(\text{SU}(N)) = N^2 - 1$ ). This gives us for  $\text{SU}(N)$ :

$$C_F = C_2(\text{fund}) = \frac{N^2 - 1}{2N}, \quad (\text{A.71})$$

$$C_A = C_2(\text{adj}) = N. \quad (\text{A.72})$$

## A.2 Frequently used relations and identities

### A.2.1 $\gamma$ -matrices

**In 4 dimensions:**

$$\gamma^\mu \gamma_\mu = 4 \quad (\text{A.73})$$

$$\{\gamma^\mu, \gamma^\nu\} = g^{\mu\nu} \quad (\text{A.74})$$

$$\gamma^\mu \gamma^\nu \gamma_\mu = -2\gamma^\nu \quad (\text{A.75})$$

$$\gamma^\mu \gamma^\nu \gamma^\rho \gamma_\mu = 4g^{\nu\rho} \quad (\text{A.76})$$

$$\gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \gamma_\mu = -2\gamma^\sigma \gamma^\rho \gamma^\nu \quad (\text{A.77})$$

$$\gamma^\nu \gamma^\rho \gamma^\sigma = g^{\mu\nu} \gamma^\rho + g^{\nu\rho} \gamma^\mu - g^{\mu\rho} \gamma^\nu - i\epsilon^{\sigma\mu\nu\rho} \gamma_\sigma \gamma^5 \quad (\text{A.78})$$

$$\gamma^\mu \not{a} \gamma_\mu = -2\not{a} \quad (\text{A.79})$$

$$\gamma^\mu \not{a} \not{b} \gamma_\mu = 4a \cdot b \quad (\text{A.80})$$

$$\gamma^\mu \not{a} \not{b} \not{c} \gamma_\mu = -2\not{c} \not{b} \not{a} \quad (\text{A.81})$$

$$\not{a} \not{b} = ab - ia_\mu \sigma^{\mu\nu} b_\nu \quad (\Rightarrow \not{a} \not{a} = a^2) \quad (\text{A.82})$$

$$\sigma^{\mu\nu} = \frac{i}{2} [\gamma^\mu, \gamma^\nu] \quad (\text{A.83})$$

**In  $D$  dimensions:**

$$g^{\mu\nu}g_{\mu\nu} = D \quad (\text{A.84})$$

$$\gamma^\mu\gamma_\mu = D \quad (\text{A.85})$$

$$\gamma^\mu\gamma^\alpha\gamma_\mu = (2 - D)\gamma^\alpha \quad (\text{A.86})$$

$$\gamma^\mu\gamma^\alpha\gamma^\beta\gamma_\mu = 4g^{\alpha\beta} - (4 - D)\gamma^\alpha\gamma^\beta \quad (\text{A.87})$$

$$\gamma^\mu\gamma^\alpha\gamma^\beta\gamma^\rho\gamma_\mu = -2\gamma^\rho\gamma^\beta\gamma^\alpha + (4 - D)\gamma^\alpha\gamma^\beta\gamma^\rho \quad (\text{A.88})$$

### A.2.2 Traces

The following formulae are still valid in  $D$  dimensions. The trace of any odd number of  $\gamma$ -matrices vanishes.

$$\text{Tr}(\gamma^\mu\gamma^\nu) = 4g^{\mu\nu} \quad (\text{A.89})$$

$$\text{Tr}(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma) = 4(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma} + g^{\mu\sigma}g^{\nu\rho}) \quad (\text{A.90})$$

$$\text{Tr}(\gamma^5) = 0 \quad (\text{A.91})$$

$$\text{Tr}(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\gamma^5) = -4i\epsilon^{\mu\nu\rho\sigma} \quad (\text{A.92})$$

$$\text{Tr}(\not{a}\not{b}) = 4ab \quad (\text{A.93})$$

$$\text{Tr}(\not{a}\not{b}\not{c}\not{d}) = 4[(a \cdot b)(c \cdot d) - (a \cdot c)(b \cdot d) + (a \cdot d)(b \cdot c)] \quad (\text{A.94})$$

$$\text{Tr}(\gamma^5\not{a}\not{b}) = 0 \quad (\text{A.95})$$

$$\text{Tr}(\gamma^5\not{a}\not{b}\not{c}\not{d}) = -4i\epsilon_{\mu\nu\rho\sigma}a^\mu b^\nu c^\rho d^\sigma \quad (\text{A.96})$$

### A.2.3 Feynman parameterization

The Feynman parameterization is used to make many loop calculations easier. For  $\frac{1}{ab}$  it can be derived in the following way:

$$\frac{1}{ab} = \frac{1}{a-b} \left( \frac{1}{b} - \frac{1}{a} \right) \quad (\text{A.97})$$

$$= \frac{1}{a-b} \int_b^a \frac{dz}{z^2}. \quad (\text{A.98})$$

We substitute:  $x = \frac{z-b}{a-b} \longrightarrow z = xa + (1-x)b, dz = (a-b)dx$

$$\implies \frac{1}{ab} = \int_0^1 dx \{a(1-x) + bx\}^{-2}. \quad (\text{A.99})$$

The case of  $\frac{1}{abc}$  (and analogously all cases with more variables) can be derived by:

$$\frac{1}{A} = \int_0^\infty dx e^{-Ax} \quad (\text{A.100})$$

$$\implies \frac{1}{abc} = \int_0^\infty dx \int_0^\infty dy \int_0^\infty dz e^{-xa-yb-zc}. \quad (\text{A.101})$$

We substitute:  $\hat{x} = x + y + z$ ,  $\hat{y} = \frac{y}{x+y+z}$ ,  $\hat{z} = \frac{z}{x+y+z}$

$$\frac{1}{abc} = \int_0^1 d\hat{y} \int_0^{1-\hat{y}} d\hat{z} \int_0^\infty d\hat{x} \hat{x}^2 \exp(-\hat{x}(\hat{y}b + \hat{z}c + (1 - \hat{y} - \hat{z})a)). \quad (\text{A.102})$$

Performing the  $\hat{x}$  integration using  $\int_0^\infty dx x^{n-1} \exp(-ax) = \frac{(n-1)!}{a^n}$  and renaming  $\hat{y} \rightarrow x, \hat{z} \rightarrow y$  gives

$$\Rightarrow \quad \frac{1}{abc} = 2 \int_0^1 dx \int_0^{1-x} dy \{a(1-x-y) + bx + cy\}^{-3}. \quad (\text{A.103})$$

#### A.2.4 Gamma- and Beta-functions

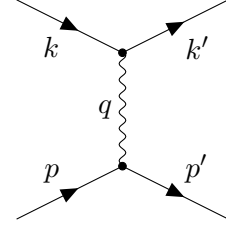
$$\Gamma(x) = \int_0^\infty z^{x-1} e^{-z} dz \quad (\text{A.104})$$

$$B(x, y) = \int_0^1 z^{x-1} (1-z)^{y-1} dz \quad (\text{A.105})$$

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \quad (\text{A.106})$$

#### A.2.5 Lab frame kinematics and kinematic variables

- $p$  : target momentum
- $M$  : target mass
- $k$  : initial momentum of incoming particle
- $k'$  : momentum of incoming particle after scattering
- $m$  : incoming particle mass
- $q$  : momentum of exchange particle
- $E$  : Energy of incoming particle
- $E'$  : Energy of incoming particle after scattering



#### Kinematic Variables

$$\nu = \frac{q \cdot p}{M} \quad (\text{A.107})$$

$$x = \frac{-q^2}{2p \cdot q} = \frac{-q^2}{2M\nu} \quad (\text{A.108})$$

$$y = \frac{p \cdot q}{p \cdot k} \quad (\text{A.109})$$



**Lab frame kinematics**

$$q^2 = -2(k \cdot k' - m^2) = 2k(k - k') = -4EE' \sin^2 \frac{\theta}{2} \quad (\text{A.110})$$

$$k \cdot k' = 2EE' \sin^2 \frac{\theta}{2} \quad (\text{A.111})$$

$$k \cdot p = EM \quad (\text{A.112})$$

$$k' \cdot p = E'M \quad (\text{A.113})$$

$$k \cdot q = k(k - k') = \frac{q^2}{2} \quad (\text{A.114})$$

$$k' \cdot q = k(k' - k) = -\frac{q^2}{2} \quad (\text{A.115})$$

$$q \cdot p = (E - E')M \quad (\text{A.116})$$

$$\nu = E - E' \quad (\text{A.117})$$

$$x = \frac{2EE'}{(E - E')M} \sin^2 \frac{\theta}{2} \quad (\text{A.118})$$

$$y = \frac{E - E'}{E} \quad (\text{A.119})$$

**A.2.6 Photon polarization sum**

In many QED calculations we get sums over the photon polarizations of the form

$$\sum_{i=1}^2 \epsilon_i^{\mu*} \epsilon_i^{\nu} \quad (\text{A.120})$$

to evaluate these sums at first we consider a spin-1 particle moving along the z-axis with momentum

$$k^\mu = E(1, 0, 0, 1) = (E, \vec{k}) \quad (\text{A.121})$$

and introduce a fixed timelike vector  $n^\mu = (1, 0, 0, 0)$ . The four polarization states are

$$\text{physical:} \quad \epsilon_1^\mu = (0, 1, 0, 0) \quad \text{and} \quad \epsilon_2^\mu = (0, 1, 0, 0), \quad (\text{A.122})$$

$$\text{unphysical:} \quad \epsilon_0^\mu = (1, 0, 0, 0) \quad \text{and} \quad \epsilon_3^\mu = (0, 0, 0, 1), \quad (\text{A.123})$$

so that  $\epsilon^\lambda \cdot \epsilon^{\lambda'} = g_{\lambda\lambda'}$ . Since the polarization vectors form an orthonormal basis in Minkowski space they obey the completeness relation

$$\epsilon_0^\mu \epsilon_0^{\nu*} - \sum_{i=1}^3 \epsilon_i^\mu \epsilon_i^{\nu*} = g^{\mu\nu}. \quad (\text{A.124})$$

Therefore we can write the sum over the physical polarizations as

$$\begin{aligned}
\sum_{i=1}^2 \epsilon_\lambda^\mu \epsilon_\lambda^{\nu*} &= -g^{\mu\nu} + \epsilon_0^\mu \epsilon_0^{\nu*} - \epsilon_3^\mu \epsilon_3^{\nu*} \\
&= -g^{\mu\nu} + n^2 - \left(\frac{k^\mu}{E} - n^\mu\right) \left(\frac{k^\nu}{E} - n^\nu\right) \\
&= -g^{\mu\nu} + n^2 - \frac{k^\mu k^\nu}{(k \cdot n)^2} + \frac{k^\mu n^\nu}{k \cdot n} + \frac{n^\mu k^\nu}{k \cdot n} - n^2.
\end{aligned} \tag{A.125}$$

Introducing  $\bar{k}^\mu = (E, -\vec{k}) = 2(k \cdot n)n^\mu - k^\mu$  we get the final result

$$\sum_{i=1}^2 \epsilon_\lambda^\mu \epsilon_\lambda^{\nu*} = -g^{\mu\nu} + \frac{k^\mu \bar{k}^\nu + k^\nu \bar{k}^\mu}{k \cdot \bar{k}}. \tag{A.126}$$

### A.3 Alternative version of figure 4.3

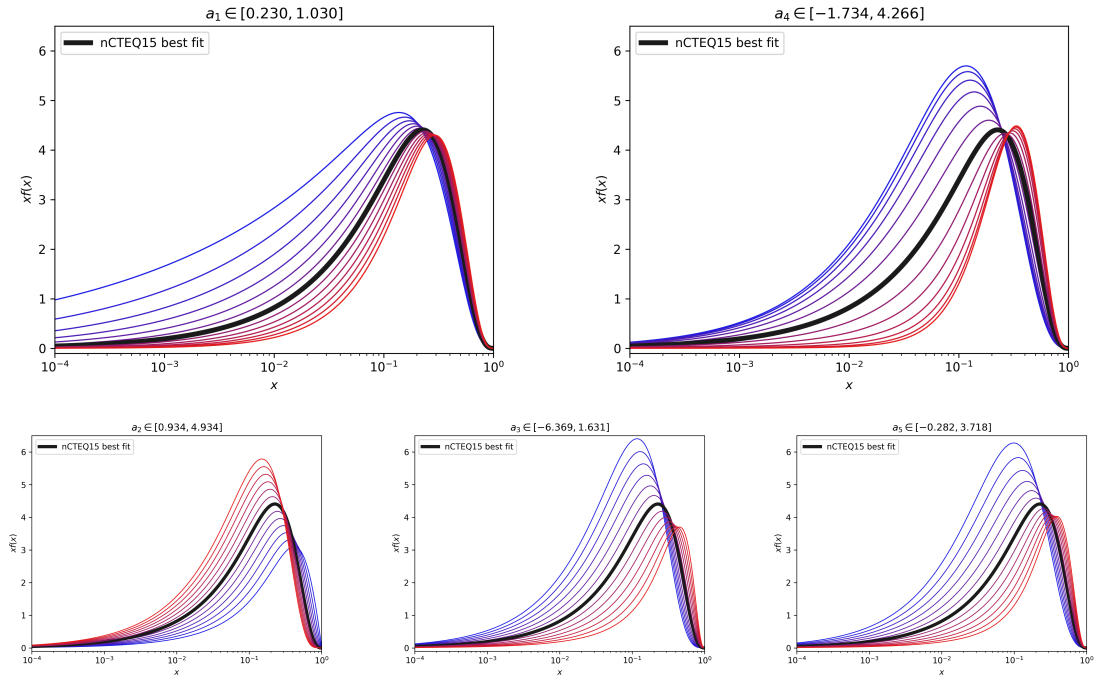


Figure A.2: Comparison of the impact of each individual parameter on the shape of the parton distribution at the input scale  $Q_0^2 = 1.69 \text{ GeV}^2$ . The blue curves represent the lower end of the given interval and the red curve the upper end. Each curve represents an equidistant step in parameter space and the solid black curve shows the best fit curve from the nCTEQ15 fit [33]. Each curve is normalized to fulfil the sum rule (88). On the top row parameters  $a_1$  and  $a_4$  are shown and smaller in the bottom row parameters  $a_2$ ,  $a_3$  and  $a_5$  because of their lower impact on the small  $x$  region.

## Bibliography

- [1] E. Rutherford. “The scattering of alpha and beta particles by matter and the structure of the atom”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 21 (1911). DOI: 10.1080/14786440508637080.
- [2] G. Aad, T. Abajyan, B. Abbott, et al. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Physics Letters B* 716 (2012). DOI: 10.1016/j.physletb.2012.08.020.
- [3] S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, et al. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Physics Letters B* 716 (2012). DOI: 10.1016/j.physletb.2012.08.021.
- [4] R. P. Feynman. *High Energy Collisions: International Conference Proceedings*. Gordon & Breach, 1970. ISBN: 978-0-677-13950-0.
- [5] M. Gell-Mann. “A schematic model of baryons and mesons”. In: *Physics Letters* 8 (1964). DOI: 10.1016/s0031-9163(64)92001-3.
- [6] Karol Kovarik, Pavel M. Nadolsky, and Davison E. Soper. “Hadron structure in high-energy collisions”. In: *Rev. Mod. Phys* (2019). arXiv: <http://arxiv.org/abs/1905.06957v1> [hep-ph].
- [7] Richard E. Taylor. *Deep inelastic scattering: the early years*. Nobel lecture. 1990. URL: <https://www.nobelprize.org/uploads/2018/06/taylor-lecture.pdf>.
- [8] Henry W. Kendall. *Deep inelastic scattering: experiments on the proton and the observation of scaling*. Nobel lecture. 1990. URL: <https://www.nobelprize.org/uploads/2018/06/kendall-lecture-1.pdf>.
- [9] Jerome I. Friedman. *Deep inelastic scattering: comparisons with the quark model*. Nobel lecture. 1990. URL: <https://www.nobelprize.org/uploads/2018/06/friedman-lecture.pdf>.
- [10] Johannes Blümlein. “The theory of deeply inelastic scattering”. In: *Progress in Particle and Nuclear Physics* 69 (2013). DOI: 10.1016/j.pnpnp.2012.09.006.
- [11] G. Altarelli and G. Parisi. “Asymptotic freedom in parton language”. In: *Nuclear Physics B* 126 (1977). DOI: 10.1016/0550-3213(77)90384-4.
- [12] Yuri L. Dokshitzer. “Calculation of the Structure Functions for Deep Inelastic Scattering and  $e^+e^-$  Annihilation by Perturbation Theory in Quantum Chromodynamics.” In: *Sov. Phys. JETP* 46 (1977).
- [13] V.N. Gribov and L.N. Lipatov. “Deep inelastic electron scattering in perturbation theory”. In: *Physics Letters B* 37 (1971). DOI: 10.1016/0370-2693(71)90576-4.
- [14] L. N. Lipatov. “The parton model and perturbation theory”. In: *Sov. J. Nucl. Phys.* 20 (1975).
- [15] Kenneth Levenberg. “A method for the solution of certain non-linear problems in least squares”. In: *Quarterly of Applied Mathematics* 2 (1944). DOI: 10.1090/qam/10666.

- [16] Donald W. Marquardt. “An Algorithm for Least-Squares Estimation of Nonlinear Parameters”. In: *Journal of the Society for Industrial and Applied Mathematics* 11 (1963). DOI: 10.1137/0111030.
- [17] J. Pumplin, D. R. Stump, and W. K. Tung. “Multivariate fitting and the error matrix in global analysis of data”. In: *Phys. Rev D* 65 (2001). DOI: 10.1103/physrevd.65.014011.
- [18] J. Pumplin, D. Stump, R. Brock, et al. “Uncertainties of predictions from parton distribution functions. II. The Hessian method”. In: *Phys. Rev D* 65 (2001). DOI: 10.1103/physrevd.65.014013.
- [19] M. Arneodo, A. Arvidson, B. Badelek, et al. “Measurement of the proton and deuteron structure functions,  $F_2p$  and  $F_2d$ , and of the ratio  $R$ ”. In: *Nucl. Phys B* 483 (1997). DOI: 10.1016/S0550-3213(96)00538-X.
- [20] A.C. Benvenuti, D. Bollini, G. Bruni, et al. “A high statistics measurement of the proton structure functions  $F_2(x, Q^2)$  and  $R$  from deep inelastic muon scattering at high  $Q^2$ ”. In: *Phys. Lett B* 223 (1989). DOI: 10.1016/0370-2693(89)91637-7.
- [21] A.C. Benvenuti, D. Bollini, G. Bruni, et al. “A high statistics measurement of the deuteron structure functions  $F_2(x, Q^2)$  and  $R$  from deep inelastic muon scattering at high  $Q^2$ ”. In: *Phys. Lett B* 237 (1990). DOI: 10.1016/0370-2693(90)91231-Y.
- [22] Joshua P. Ellis. “TikZ-Feynman: Feynman diagrams with TikZ”. In: *Comput. Phys. Commun* 210 (2017). DOI: 10.1016/j.cpc.2016.08.019.
- [23] F. Halzen and A. Martin. *Quarks and Leptons*. John Wiley & Sons, Jan. 6, 1984. 416 pp. ISBN: 0471887412.
- [24] S. Moch, J.A.M. Vermaseren, and A. Vogt. “The longitudinal structure function at the third order”. In: *Physics Letters B* 606 (2005). DOI: 10.1016/j.physletb.2004.11.063.
- [25] J. D. Bjorken. “Current algebra at small distances”. In: *Conf. Proc.* C670717 (1967).
- [26] J. D. Bjorken. “Asymptotic Sum Rules at Infinite Momentum”. In: *Physical Review* 179 (1969). DOI: 10.1103/physrev.179.1547.
- [27] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Pr., 2014. ISBN: 1107034736.
- [28] Daniel V. Schroeder Michael E. Peskin. *An Introduction To Quantum Field Theory*. Taylor & Francis Inc, Sept. 11, 1995. 866 pp. ISBN: 0201503972.
- [29] G. Curci, W. Furmanski, and R. Petronzio. “Evolution of parton densities beyond leading order”. In: *Nuclear Physics B* 175 (1980). DOI: 10.1016/0550-3213(80)90003-6.
- [30] W. Furmanski and R. Petronzio. “Singlet parton densities beyond leading order”. In: *Physics Letters B* 97 (1980). DOI: 10.1016/0370-2693(80)90636-X.
- [31] S. Moch, J.A.M. Vermaseren, and A. Vogt. “The three-loop splitting functions in QCD: the non-singlet case”. In: *Nuclear Physics B* 688 (2004). DOI: 10.1016/j.nuclphysb.2004.03.030.

- [32] A. Vogt, S. Moch, and J.A.M. Vermaseren. “The three-loop splitting functions in QCD: the singlet case”. In: *Nuclear Physics B* 691 (2004). DOI: 10.1016/j.nuclphysb.2004.04.024.
- [33] K. Kovarik, A. Kusina, T. Jezo, et al. “nCTEQ15 - Global analysis of nuclear parton distributions with uncertainties in the CTEQ framework”. In: *Phys. Rev. D* 93, 085037 (2016). DOI: 10.1103/PhysRevD.93.085037.
- [34] G.P. Salam and J. Rojo. “A Higher Order Perturbative Parton Evolution Toolkit (HOPPET)”. In: *Computer Physics Communications* 180 (2009). DOI: 10.1016/j.cpc.2008.08.010.
- [35] D. Stump, J. Pumplin, R. Brock, et al. “Uncertainties of predictions from parton distribution functions. I. The Lagrange multiplier method”. In: *Phys. Rev D* 65 (2001). DOI: 10.1103/physrevd.65.014012.
- [36] Conrad Sanderson and Ryan Curtin. “Armadillo: a template-based C++ library for linear algebra”. In: *The Journal of Open Source Software* 1 (2016). DOI: 10.21105/joss.00026.
- [37] Brian P. Flannery, Saul Teukolsky, William H. Press, et al. *Numerical recipes in C++*. Cambridge University Press, 1986.
- [38] Andy Buckley, James Ferrando, Stephen Lloyd, et al. “LHAPDF6: parton density access in the LHC precision era”. In: *The European Physical Journal C* 75 (2015). DOI: 10.1140/epjc/s10052-015-3318-8.
- [39] Tancredi Carli, Dan Clements, Amanda Cooper-Sarkar, et al. “A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: the APPLGRID project”. In: *The European Physical Journal C* 66 (2010). DOI: 10.1140/epjc/s10052-010-1255-0.
- [40] I. Antcheva, M. Ballintijn, B. Bellenot, et al. “ROOT A C++ framework for petabyte data storage, statistical analysis and visualization”. In: *Computer Physics Communications* 180 (2009). DOI: 10.1016/j.cpc.2009.08.005.
- [41] Pavel M. Nadolsky, Hung-Liang Lai, Qing-Hong Cao, et al. “Implications of CTEQ global analysis for collider observables”. In: *Phys. Rev D* 78 (2008). DOI: 10.1103/physrevd.78.013004.
- [42] K. Kovarik. “Unpublished manuscript”. "Hitchhiker’s guide to renormalization".