IEKP-KA/2013-10

Karlsruhe Institute of Technology

Bestimmung von Protonstruktur und starker Kopplung aus inklusiven Jet-Wirkungsquerschnitten am LHC

Georg Sieber

Diplomarbeit

an der Fakultät für Physik
des Karlsruher Instituts für Technologie (KIT)

Referent: Prof. Dr. G. Quast
Institut für Experimentelle Kernphysik

Korreferent: Prof. Dr. U. Husemann
Institut für Experimentelle Kernphysik

15. März 2013

Deutsche Zusammenfassung

Sensitivität der Partonverteilungsfunktionen des Protons auf inklusive Jet-Wirkungsquerschnitte von CMS

Mitte des letzten Jahrhunderts wurde das Standardmodell der Teilchenphysik entwickelt. Begünstigt durch systematische Experimente und neue theoretische Modelle war es erstmals möglich, alle beobachteten Phänomene innerhalb eines Modells zu erklären. Die moderner Teilchenphysik kennt vier fundamentale Wechselwirkungen: Die Gravitationskraft, die elektromagnetische Kraft, die schwache Kraft sowie die starke Kraft. Alle diese Grundkräfte bis auf die Gravitationskraft sind einheitlich im Standardmodell beschrieben. Das Modell beinhaltet jedoch einige freie Parameter, die in Experimenten bestimmt werden müssen. Da einige dieser Parameter wie z.B. die starke Kopplung energieabhängig sind, müssen sie in den neu zugänglichen Phasenraumbereichen an modernen Teilchenbeschleunigern überprüft werden. Des weiteren ermöglicht die Erforschung dieser unbekannten Phasenraumbereiche die Suche nach Physik jenseits des Standardmodells, sowie die Überprüfung von Vorhersagen des Standardmodells bei hohen Energien.

Aus diesem Grund wurde der Large Hadron Collider (LHC) am Forschungszentrum CERN in der Nähe von Genf gebaut. In unterirdischen Tunneln werden gegenläufige Protonstrahlen zur Kollision gebracht und die Wechselwirkungen untersucht. Hierzu befinden sich vier Großexperimente an den Kollisionspunkten, die diese vermessen und aufzeichnen. Eines dieser Experimente ist der Compact Muon Solenoid (CMS). Dieser Detektor ist dafür ausgelegt, einen großen Phasenraumbereich untersuchen zu können. Ein besonderer Schwerpunkt wurde hierbei auf die Suche nach dem vom Standardmodell vorhergesagten Higgs-Boson gelegt. Im Juli 2012 verkündeten die ATLAS- und CMS-Kollaborationen die unabhängige Entdeckung eines neuen Teilchens, das konsistent mit dem vorhergesagten Higgs-Boson ist.

Da das Proton kein elementares Teilchen ist, sondern eine Substruktur besitzt, werden die Kollisionen der Konstituenten des Protons beobachtet. Für die theoretische Vorhersage der möglichen Prozesse ist daher ein besonders genaues Wissen über die Substruktur des Proton wichtig. In vielen Studien am LHC bei hohen Energien stellen die Grenzen dieses Wissens eine der dominierenden Unsicherheitsquellen dar. Das Proton lässt sich über Partonverteilungsfunktionen (PDFs) beschreiben, die den Impulsbruchteil der jeweiligen Partonen angeben. Diese PDFs können nicht störungstheoretisch auf Basis der Quantenchromodynamik (QCD)

berechnet werden, sondern müssen aus experimentellen Messungen abgeleitet werden.

Bei der Kollision von zwei Konstituenten des Protons können bei genügend hohen Energien *Jets* entstehen. Bei der Entstehung der Jets werden neue Quark-Antiquark Paare erzeugt, deren Flugrichtung kollinear zu der des ursprünglichen Teilchens ist. Deshalb können sie als Jets zusammengefasst werden. Die Eigenschaften der Jets lassen hierbei Rückschlüsse auf das ursprüngliche Teilchen und somit auf die Struktur des Protons zu. Die Produktion von hadronischen Jets ist einer der dominierenden Prozesse am LHC. Mit Hilfe des CMS-Detektors wurde der inklusive Jet-Wirkungsquerschnitt bei einer Schwerpunktsenergie von 7 TeV gemessen. Das Spektrum des Transversalimpulses wurde getrennt für die Rapidität der Jets analysiert, da bestimmte Phasenraumbereiche der PDFs auf unterschiedliche Rapiditäten der Jets sensitiv sind. Mit Hilfe dieser Messung können Vorhersagen der QCD bei hohen Energien überprüft werden und letztendlich die PDFs genauer bestimmt werden. Die theoretischen Vorhersagen werden auf der Basis der Quantenchromodynamik (QCD) störungstheoretisch in nächstführender Ordnung berechnet.

Die kürzlich veröffentlichte Software HERAFitter wird eingesetzt um die PDFs des Protons zu bestimmen. HERAFitter ist ein von der H1 und ZEUS Kollaboration entwickeltes Framework, das es ermöglicht aus Wirkungsquerschnitts-Messungen verschiedener Observablen diese PDFs mit den Fehler abzuschätzen. Da insbesondere die PDF des Gluons nur unzureichend mit Daten aus tiefinelastischer Streuung beschrieben werden kann, bieten Messungen des inklusiven Jet-Wirkungsquerschnitts des CMS Experiments einen wertvollen Beitrag zur Reduzierung der Unsicherheit. Gleichzeitig kann die Kompatibilität mehrerer Messungen von verschiedenen Experimenten in einem gemeinsamen Fit überprüft und systematische Unsicherheiten studiert werden. Zur Überprüfung des Einflusses der inklusiven Jet-Wirkungsquerschnitte wurden Fits der PDFs mit diesen Daten durchgeführt und mit PDFs verglichen, die ohne diese Daten bestimmt wurden. Hierbei wurden die einzelnen Unsicherheitsquellen sehr genau studiert und separat betrachtet: Die Unsicherheiten der Daten wurden auf die PDFs fortgepflanzt und sind in dem roten Unsicherheitsband zu finden. Unsicherheiten, die auf Modellparameter wie die Quarkmassen zurückgehen, sind in dem gelben Unsicherheitsband gebündelt. Da die Fits der PDFs eine initiale Parametrisierung erfordern, die willkürlich gewählt wird, wurden die Einflüsse weiterer möglicher Parametrisierungen studiert und in dem grünen Band zusammengefasst.

Das Ziel dieser Studie ist, die PDFs zu verbessern und die Fehler zu reduzieren. Dies kann am Beispiel der Gluon-PDF in Abbildung 1 sehr gut beobachtet werden. Die Gluon-PDF ist besonders sensitiv auf die inklusiven Jet-Daten von CMS, wodurch eine deutliche Reduzierung der Unsicherheiten sichtbar wird.

Auch die Quark PDFs können mit Hilfe der inklusiven Jet-Daten verbessert werden. Insbesondere bei sehr hohen Transversalimpulsen und im Vorwärtsbereich ist der Anteil an qq Prozessen in der Jet-Produktion dominant und somit die Sensiti-

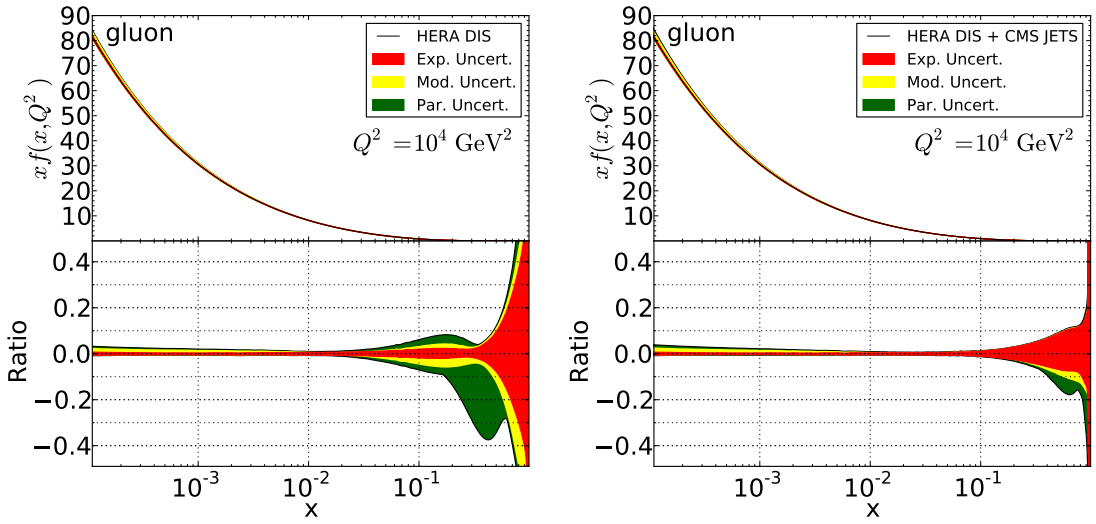


Abbildung 1.: PDF-Fit des Gluons mit tiefinelastischen Daten des HERA-Beschleunigers (links) und zusätzlich mit inklusiven Jet-Wirkungsquerschnitten von CMS (rechts). Die PDF des Gluons ist direkt zugänglich unter Zuhilfenahme der Jet-Wirkungsquerschnitte. Dies führt zu einer deutlichen Reduzierung der Unsicherheitsquellen der PDFs.

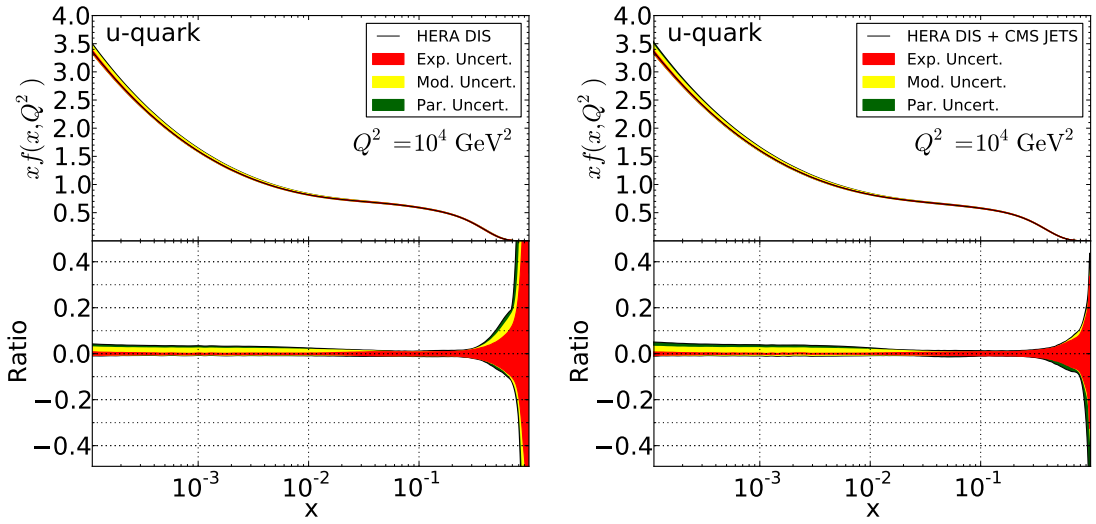


Abbildung 2.: PDF-Fit des Up-Quarks. *qq* Prozesse in der inklusiven Jet Produktion lassen Rückschlüsse auf die Quark-PDFs zu und führen zu einer weiteren Reduzierung der Unsicherheiten.

vität der Quark-PDFs hoch. Anhand des Vergleichs der PDFs in Abbildung 2 sieht man, dass die Unsicherheit auch hier verringert werden konnte.

Bestimmung der starken Kopplung aus inklusiven Jet-Wirkungsquerschnitten von CMS

Die Sensitivität der starken Kopplung auf inklusive Jet-Wirkungsquerschnitte ermöglicht die Bestimmung derselben aus den Daten. Hierzu werden PDFs eingesetzt, die für verschiedene Werte von α_s bestimmt wurden. Aus Vergleichen der Vorhersagen mit jedem dieser PDF kann die starke Kopplungskonstante bestimmt werden. Die NNPDF-Kollaboration veröffentlicht diese PDF-Sets mit Werten der starken Kopplung $\alpha_s(M_Z^2)$ von 0.106 bis 0.124 in Schritten von 0.001. Abbildung 3 zeigt die Vorhersagen des Jet-Wirkungsquerschnitts als Verhältnis zu der Vorhersage für $\alpha_s(M_Z^2) = 0.119$. Die Datenpunkte mit ihren totalen Unsicherheitsquellen werden in diesem Bereich gut beschrieben und lassen somit die Bestimmung der starken Kopplungskonstanten zu.

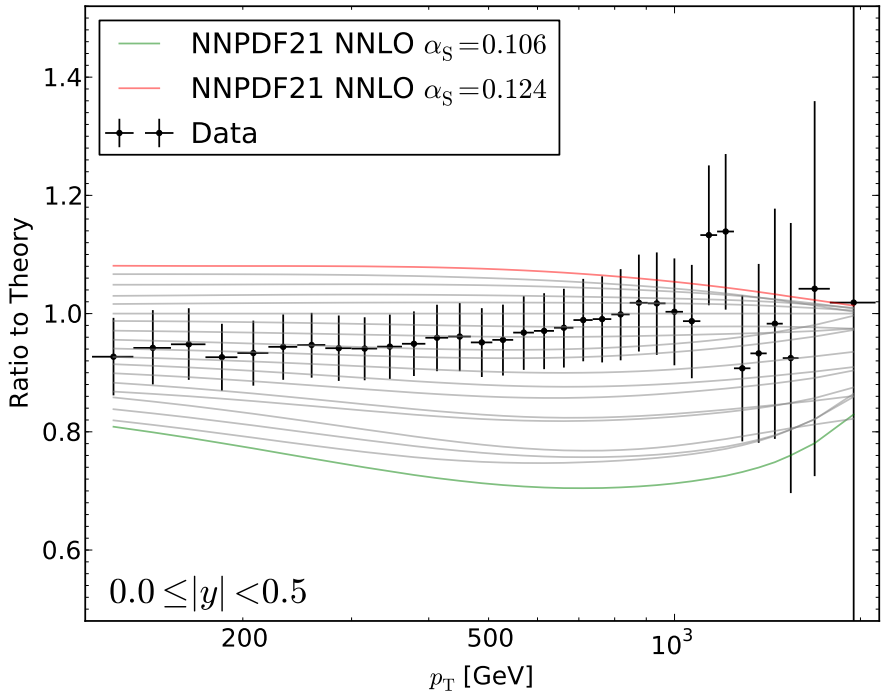


Abbildung 3.: NNPDF veröffentlicht PDF Fits mit Werten der starken Kopplung $\alpha_s(M_Z^2)$ von 0.106 bis 0.124 in Schritten von 0.001. Die Vohersagen mit diesen PDFs werden im Verhältnis zur Vorhersage $\alpha_s(M_Z^2) = 0.119$ dargestellt. Der Vergleich mit den Datenpunkten ergibt eine gute Übereinstimmung der Vohersagen mit den gemessenen Wirkungsquerschnitten.

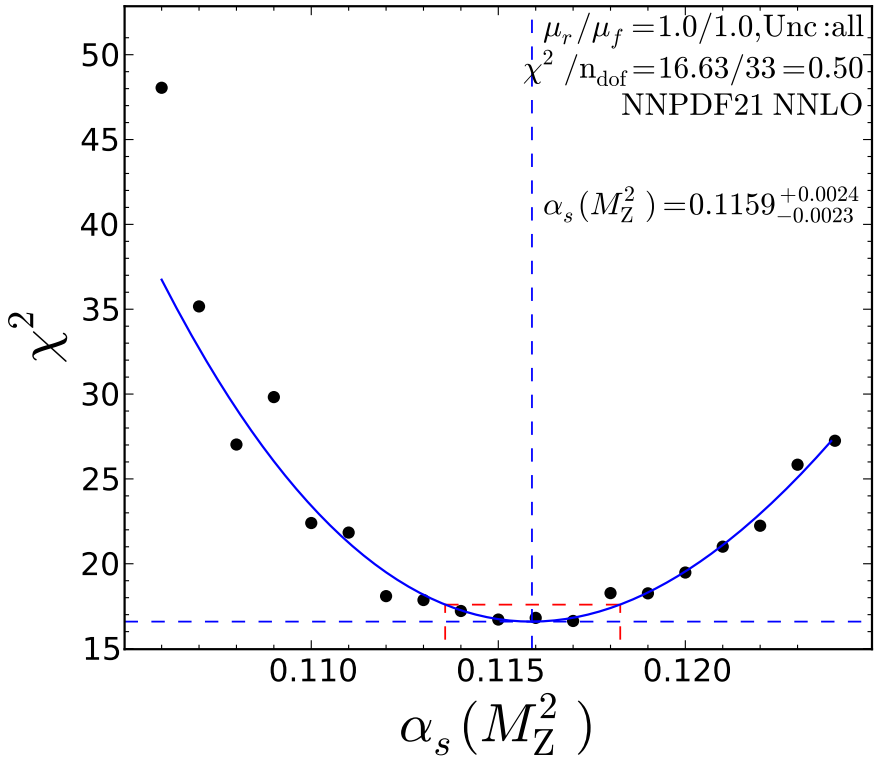


Abbildung 4.: Die gemessenen Daten werden mit Hilfe eines χ^2-Tests mit den Theorievorhersagen verglichen und der beste Wert für $\alpha_s(M_Z^2)$ wird bestimmt. Hierbei werden alle experimentellen und theoretischen Unsicherheitsquellen bis auf die Skalenunsicherheit berücksichtigt, die separat behandelt wird. Der rot gestrichelte Bereiche markiert das 68% Konfidenzintervall.

Mit Hilfe eines χ^2-Tests wurde die Übereinstimmung der Daten mit den Vorhersagen unter Berücksichtigung aller Fehlerquellen überprüft und der optimale Wert von $\alpha_s(M_Z^2)$ bestimmt. In Abbildung 4 ist die χ^2-Verteilung für den zentralen Rapiditätsbereich des inklusiven Jet-Wirkungsquerschnitts dargestellt. Die rot gestrichelten Linien markieren das 68%-Konfidenzintervall. Des weiteren wurden die experimentellen und theoretischen Fehler abgeschätzt und das Ergebnis mit anderen Analysen verglichen.

$$\alpha_s(M_Z^2) = 0.1159\ ^{+0.0016}_{-0.0015}(\text{exp})\ ^{+0.0018}_{-0.0017}(\text{PDF})\ ^{+0.0018}_{-0.0009}(\text{Skala})$$

Hierbei wurde die Übereinstimmung mit ähnlichen Analysen von CMS und DØ festgestellt und auch der Weltmittelwert von $\alpha_s(M_Z^2)$ ist innerhalb der Fehlerquellen kompatibel mit dem Resultat dieser Studie.

Die hohe Schwerpunktenergie erlaubt, das Spektrum des inklusiven Jet-Wirkungsquerschnitts bis zu Transversalimpulsen von 2100 GeV im zentralen Rapiditätsbereich zu messen. Hierzu wird der zentrale Rapiditätsbereich in die drei Berei-

che 114 GeV bis 300 GeV, 300 GeV bis 737 GeV und 737 GeV bis 2116 GeV unterteilt und die starke Kopplung separat in jedem Bereich extrahiert. Die Energieabhängigkeit der starken Kopplung wird von der QCD vorhersagt und kann mit Hilfe der DGLAP-Evolutionsgleichungen bestimmt werden. Abbildung 5 zeigt die Vorhersage der laufenden Kopplung (schwarz) sowie mehrere Messungen, die mit dem Laufen der Kopplung in Energiebereichen von 120 GeV bis 800 GeV übereinstimmen.

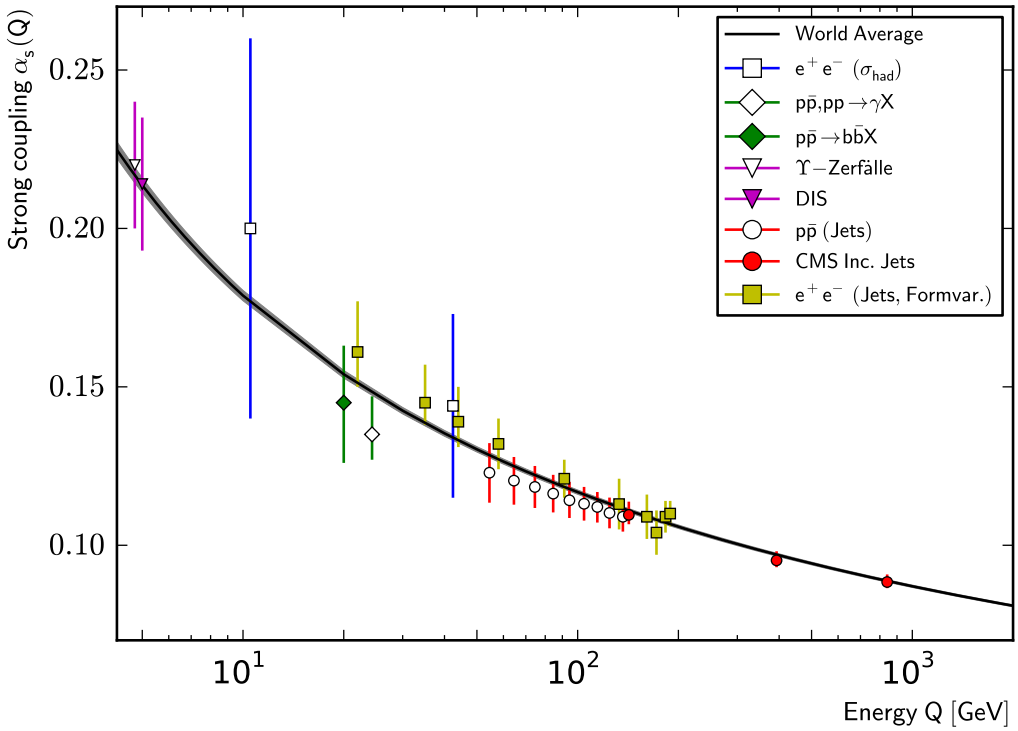


Abbildung 5.: Überblick über verschiedene Messungen der starken Kopplungskonstanten, die die Energieabhängigkeit der starken Kopplung bestätigen. Zusätzlich wurde die eigene Messung (rote Punkte) hinzugefügt, die gut mit der vorhergesagten Energieabhängigkeit übereinstimmen.

Die Extraktion der starken Kopplungskonstanten aus inklusiven Jet-Wirkungsquerschnitten ergibt 3 Werte für die starke Koppplung in unterschiedlichen Energiebereichen. Diese können mit dem vorhergesagtem Laufen der Kopplung verglichen werden (rote Punkte) und sind innerhalb der Fehler kompatibel.

Karlsruhe Institute of Technology

Determination of the proton structure and the strong coupling from inclusive jet cross sections at the LHC

Georg Sieber

Diplomarbeit

an der Fakultät für Physik
des Karlsruher Instituts für Technologie (KIT)

Referent: Prof. Dr. G. Quast
Institut für Experimentelle Kernphysik

Korreferent: Prof. Dr. U. Husemann
Institut für Experimentelle Kernphysik

March 15th, 2013

Contents

1. Introduction

Physicists at the LHC reproduce conditions comparable to the time shortly after the Big Bang by colliding particles at high energies. A deeper understanding of the phenomena in high-energy physics will provide insights into the evolution of the early universe towards the current state.

Present knowledge about the fundamental particles and their interactions is summarised within the Standard Model of particle physics. While it is incredibly successful in describing most of the observed phenomena, several issues remain unresolved, as the model relies on numerous parameters which have to be determined by experiments. The most recent success was the discovery of a Higgs-like boson in the mass range predicted by the Standard Model.

The Large Hadron Collider (LHC) is the latest step in reaching a deeper understanding of the early state of the universe. Proton-proton beams are brought to collision at a centre-of-mass energy never reached before. Since the proton is not an elementary particle, the collisions of its constituents, quarks and gluons, are observed. A detailed understanding of the proton's substructure is necessary for precision studies evaluating predictions of the Standard Model. More profound information about the proton will reduce one of the dominant uncertainty sources in current analyses at the LHC.

Interactions of the proton's constituents produce high-energetic particles. These are registered as a collimated spray of particles in the detector. Information about the properties of the basic constituents of matter and the strong interaction is gained by analysing jets.

The LHC enables probing the Standard Model at energies never reached before. Jet production is the dominant process at hadron colliders and an ideal candidate to study the Standard Model's predictions at high energies. The inclusive jet cross section is measured differentially in the transverse momentum of the jet and the rapidity.

The momentum fraction carried by the proton's constituents is described by the so-called parton distribution functions (PDFs). Since the inclusive jet cross section is dependent on these PDFs in certain phase-space regions, the PDFs themselves can be constrained by the inclusive jet measurement. The theory predictions needed for comparisons with the data are performed on the basis of quantum chromodynamics (QCD) in which the observable is calculated in a perturbative approach for scattering processes with a large momentum transfer.

The determination of the proton structure is performed in a combined fit to data from various collider experiments within the framework of HERAFitter. HERAFitter is a recently published software tool by the H1 and ZEUS collaborations providing the possibility to include own data in PDF fits. Users can study the constraints of their measurements on the PDFs and their uncertainties. The main constraints on the PDFs derive from measurements done at HERA[1]. Additional measurements from the LHC constrain the PDFs in phase space regions not reachable by HERA measurements. The inclusive jet cross section will help to constrain the PDFs in these regions.

The dependence of the inclusive jet cross section on the value of the strong coupling constant $\alpha_s(M_Z^2)$ makes it possible to extract α_s either simultaneously with or separately from the determination of the PDFs.

Within this thesis, the sensitivity of the PDFs and the strong coupling to the inclusive jet cross section from the CMS experiment has been studied with the aim to further constrain the PDFs and gain a better understanding of the proton. Additionally the strong coupling will be determined at scales close to the TeV range.

[1] Hadron-Elektron-Ringanlage

2. Theory of the Strong Interaction

The majority of the visible universe is built up by baryonic matter, held together by the strong interaction. This interaction is responsible for the binding of quarks and anti-quarks to hadrons at small scales. It also mediates the binding force of protons and neutrons within the nucleus of an atom. The strong interaction is described within the Standard Model by the theory of quantum chromodynamics (QCD) and has been very successful in predicting the interaction between quarks at very high momentum transfers, which correspond to very small distances. Several techniques have been developed which ease the often very complicated QCD calculations. At high energies, the theory allows perturbative QCD (pQCD) calculations. This is currently the most precise approach to QCD with a reasonable calculation time. Lattice QCD approaches the strong interaction using a discrete set of space-time points to reduce the very complicated path integrals to numerical computations which can be solved on supercomputers. It has some advantages over pQCD as it is not limited to high transverse momenta, but the problems of this very resource-intensive approach as well as numerical problems lead to the preference of perturbative methods.

2.1. Standard Model of Particle Physics

In the mid of the 20th century, the Standard Model has been developed in a collaborative approach by experimentalists and theorists. Experimental discoveries and new theoretical models were combined into a comprehensive model able to describe the observed phenomena. Since then, the Standard Model has been probed in a large number of different experiments and it showed an impressive agreement between the predictions and experimental results.

The Standard Model contains 12 fermions, spin-$\frac{1}{2}$ particles, from which the known matter in the universe is built up. Each of these fundamental particles has its own anti-particle with opposite quantum numbers. These particles are ordered in three so-called generations according to their properties.

The twelve fermions interact with each other through four fundamental forces, the weak, the strong, the electromagnetic and the gravitational force, of which the first three are described by the Standard Model. As gravity is too weak to have an impact on physics on microscopic scales, this poses no problem for high-energy physics. Step-by-step all of the predicted particles were observed and the confidence in the Standard Model has increased with each confirmation.

fermions	generation 1	2	3	charge	weak isospin	colour
quarks	u	c	t	$+\frac{2}{3}$	$+\frac{1}{2}$	r,g,b
	d'	s'	b'	$-\frac{1}{3}$	$-\frac{1}{2}$	
leptons	ν_e	ν_μ	ν_τ	0	$+\frac{1}{2}$	-
	e	μ	τ	-1	$-\frac{1}{2}$	

Table 2.1.: The twelve fermions described within the Standard Model. For each fermion, the charges, the third component of the weak isospin and the colour charge is shown. The corresponding anti-particles have opposite quantum numbers. The marker for the down-type quarks indicates the weak eigenstates of the quarks which can be transformed into mass eigenstates via the CKM[1]-matrix.

particle	interaction	mass [GeV]	J^P	q	T_3
photon γ	elect.mag.	0	1^-	0	0
Z^0	weak	91.18	1	0	0
$W^\pm$	weak	80.40	1	$\pm$e	±1
gluons	strong	0	1^-	0	0

Table 2.2.: The properties of the gauge bosons mediating the four fundamental interaction forces in the Standard Model: The mass, the spin and parity, the charge and the weak isospin.

However, the Standard Model is not an all-explaining theory, as there are many free parameters which need to be determined by experiments.

- the masses of the leptons and the quarks
- the mass of the W/Z boson and of the Higgs boson
- the coupling of the electromagnetic interaction, α, and the strong coupling α_s
- the four parameters of the weak mixing matrix (CKM) of the quarks and leptons describing the transition of mass eigenstates to weak eigenstates.

Furthermore, there are experimental discoveries, which cannot be explained within the Standard Model like oscillation of the neutrino flavours. The neutrinos are massless in the Standard Model, but the observed oscillation can only be explained if neutrinos are massive. The transition probabilities are summarised in the PMNS[2] matrix. Another outstanding issue of the Standard Model was the predicted, but not yet found Higgs boson and its mass. On July 4, 2012, CMS and

[2]Pontecorvo-Maki-Nakagawa-Sakata

ATLAS both announced the discovery of a Higgs-like boson, which agrees with the expected properties of a Higgs boson, at a mass of around 125 GeV [1, 2].

2.2. Quantum Chromodynamics

Quantum chromodynamics is the theory of the strong interaction, one of the four fundamental forces. QCD is a quantum field theory consisting of colour fields mediated by gluons. It describes the interactions between quarks and gluons and is responsible for the creation of hadrons and the binding of protons and neutrons into nuclei as well.

The observation of bound quark states consisting of three up-quarks with parallel spin (Δ^{++}) seemed to violate the Pauli Principle, which expects the particle wave function to be anti-symmetric. To restore the anti-symmetry of the wave function, an additional quantum number has been introduced. This quantum number is called colour-charge and the name of the charges are red, green and blue with their contrary anti-colour charges. The additional quantum number leads to a anti-symmetric wave function and explains the observations in consistency with the Pauli Principle.

symmetry	colour states							
octet	$r\bar{g}$	$r\bar{b}$	$g\bar{b}$	$g\bar{r}$	$b\bar{r}$	$b\bar{g}$	$\sqrt{1/2}(r\bar{r}-g\bar{g})$	$\sqrt{1/6}(r\bar{r}+g\bar{g}-2b\bar{b})$
singlet	$\sqrt{1/3}(r\bar{r}+g\bar{g}+b\bar{b})$							

Table 2.3.: One possible representation of the gluon colour-states.

2.2.1. The QCD Lagrangian

The theory of the interaction between quarks and gluons was initially formulated by Fritzsch, Gross, Wilczek and Weinberg as a quantum field theory, more specifically a non-abelian gauge theory [3, 4, 5]. The dynamics of the quantum fields can be derived from the action using basic field theory. The action is defined in terms of the Lagrangian density as

$$S = i \int d^4x \mathcal{L}(x) \tag{2.1}$$

The equation of motion is found from the the minimisation of S using the Euler-Lagrange equation.

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi_i)} \right) - \frac{\partial \mathcal{L}}{\partial \psi_i} = 0 \tag{2.2}$$

The Lagrangian for spin-$\frac{1}{2}$ spinor field $\psi(x)$ results from quantum mechanics.

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \tag{2.3}$$

The Euler-Lagrange-equation applied to this Lagrangian gives the Dirac-equation.

$$(i\gamma^\mu \partial_m u - m)\psi(x) = 0 \tag{2.4}$$

The equation of motion for a Lagrangian containing normal derivatives is not invariant under transformations between possible gauges. Since this would break the gauge invariance, the minimal substitution method is applied which replaces the derivative ∂_μ with D_μ . This derivative introduces eight additional gauge fields A^a_μ to make the Lagrangian invariant under gauge transformations.

$$\partial_\mu \rightarrow D_\mu := \partial_\mu - igA^a_\mu T^a \tag{2.5}$$

The additional term in the Lagrangian introduces the interaction between the gauge field A^a_μ and the field ψ.

$$\mathcal{L} = \bar{\psi}(i\gamma_\mu \partial\mu - m)\psi + g_s \bar{\psi}\gamma_\mu A^a_\mu T^a \psi \tag{2.6}$$

To fully describe the gauge fields A^a_μ, derivative terms have to be introduced which keep gauge and Lorentz invariance. In fact this can be done by the introduction of a gauge field strength tensor $F^a_{\mu\nu}$.

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_s f^{abc} A^b_\mu A^c_\nu \tag{2.7}$$

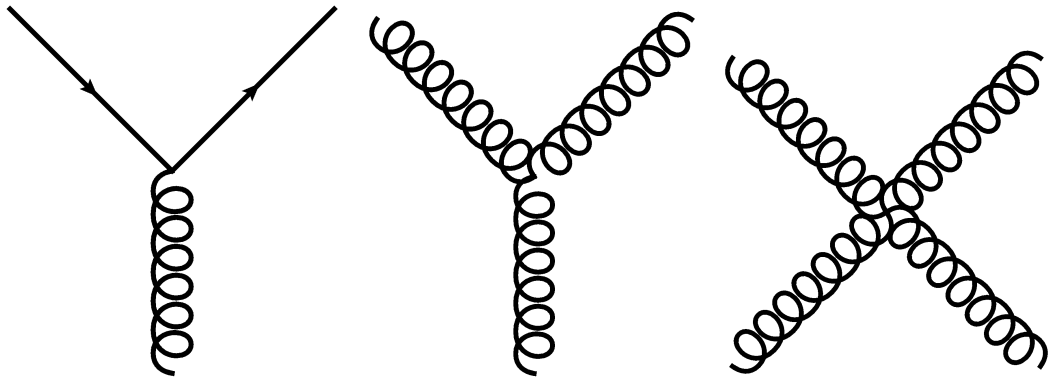
This additional Lagrangian term is called the gauge part.

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu} \tag{2.8}$$

Free parameters of this theory are the gauge coupling g_s and the quark masses. The strong coupling directly relates to the gauge coupling via $\alpha_\text{s} = g_s^2/4_\pi$. The fundamental couplings of the gluon are shown in Figure 2.1. Since QCD is a non-abelian gauge theory, the self-coupling of the colour-carrying gluons leads to additional couplings compared to QED.

2.2.2. The Running Coupling of the Strong Interaction

In analogy to quantum electrodynamics (QED) with the fine-structure constant α, a coupling constant α_s has been introduced for QCD. The strong coupling constant describes the strength of the interaction between the gluon and the colour charge $g_s = \sqrt{4\pi\alpha_\text{s}}$. Just as in QED, the coupling constant is not energy-independent, but changes with the squared momentum transfer Q^2. The energy dependence is due to particles originating from vacuum fluctuations, which interact with the electrical charge or colour-charge. These interactions should lead to an increase of the strong coupling constant at higher Q^2. But in QCD the gluons couple to



(a) (b) (c)

Figure 2.1.: The fundamental couplings of the gluons:a) a quark interacting with a gluon and b,c) the gluon-gluon self-coupling.

themselves, leading to a dominating anti-screening effect. Therefore, effectively α_s decreases at small distances (i.e. high momentum transfers).

In processes with high momentum transfer, the coupling gets very small and the quarks can be regarded as free. This is also called *asymptotic freedom*. However, if the momentum transfer is very small, the coupling can become very strong and perturbative techniques are not applicable anymore. Therefore all colour charges are bound together into colourless objects, and a separation of two coloured particles is not possible, but leads to the energetically more favourable creation of new colourless hadrons. This phenomenon is named *confinement*.

Energy Dependence of α_s

Physical observables can be written in a perturbative approach as an expansion in the strong coupling α_s. Since it is possible that additional emissions and ultraviolet divergences arise, these have to be removed in a process called renormalisation. This introduces an additional energy scale μ, and α_s and physical observables get dependent on μ. The fact that physical observables X do not depend on these scale is given mathematically by the renormalisation group equation (RGE):

$$\mu^2 \frac{d}{d\mu^2} X(\frac{q^2}{\mu^2}, \alpha_s(\mu^2)) \stackrel{!}{=} 0 = \left(\mu^2 \frac{\partial}{\partial \mu^2} + \mu^2 \frac{\partial \alpha_s(\mu^2)}{\partial \mu^2} \frac{\partial}{\partial \alpha_s(\mu^2)} \right) X(\frac{q^2}{\mu^2}, \alpha_s(\mu^2) \quad (2.9)$$

The behaviour of the strong coupling constant can be explained by the introduction of a β-function

$$\frac{\partial \ln \alpha_s(\mu^2)}{\partial \ln \mu^2} = \frac{\beta(\alpha_s(\mu^2))}{\alpha_s(\mu^2)} \quad (2.10)$$

A perturbative ansatz for the solution of the β function is

$$\beta(\alpha_s(q^2)) = -\frac{\beta_0}{4\pi}\alpha_s^2(q^2) - \frac{\beta_1}{8\pi^2}\alpha_s^3(q^2) + \mathcal{O}(\alpha_s^4) \tag{2.11}$$

The coefficients β_i can be calculated within a given renormalisation scheme like the $\overline{MS}$ [6] scheme. With a one-loop QCD calculation, and a given number of quark flavours n_f, β_0 can be expressed as

$$\beta_0 = \frac{33 - 2n_f}{12\pi} \tag{2.12}$$

$$\beta_1 = \frac{153 - 19n_f}{24\pi^2} \tag{2.13}$$

Solving the RGE equation for the one-loop solution of the β function leads to

$$\alpha_s(q^2) = \frac{\alpha_s(\mu^2)}{1 + \alpha_s(\mu^2)\frac{\beta_0}{4\pi}\ln(\frac{q^2}{\mu^2})}. \tag{2.14}$$

To compare the strong coupling constant between different experiments, it is common usage to evaluate α_s at $\mu^2 = (M_Z)^2$. The results, which are available at next-to-next-to-leading-order (NNLO) precision, are combined into a world average value of α_s. The Particle Data Group reports α_s currently as

$$\alpha_s(M_Z^2) = 0.1184 \pm 0.0007. \tag{2.15}$$

2.2.3. Hadrons

The confinement ensures the colour neutral state of observable objects. Since separating quarks from each other needs more energy than the creation of another quark-antiquark pair, the quarks combine into colourless bound states. These are called hadrons and can be classified according to their constituents.

Mesons consist of one quark and one anti-quark. They have spin-0 or spin-1 and are therefore bosons.

Baryons consist of three quarks or anti-quarks. They have spin-$\frac{1}{2}$ and are fermions.

All hadrons except for the proton are unstable. The decay of a proton is studied, but has not yet been observed experimentally. The decay of such hadrons occur via the strong, weak or electro-magnetic interaction.

The properties of the proton have been studied very carefully at particle colliders. The ideal way to probe the substructure of the proton is looking at deep inelastic scattering (DIS) of electrons or positrons through the exchange of a virtual photon. Above an energy of around 1 GeV, the inner substructure of the proton is revealed and it was found that the proton consists of point-like particles, the so-called partons. The behaviour of the substructure of the proton can be described very well in the parton-model and even better in the QCD improved parton model.

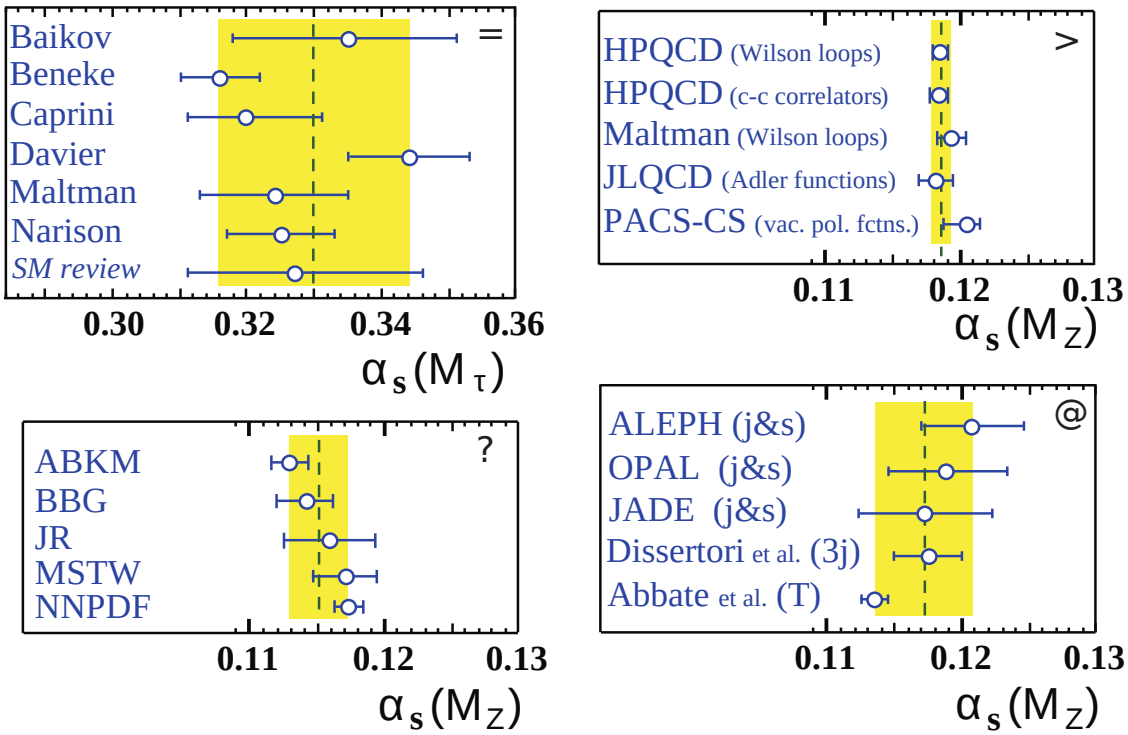


Figure 2.2.: Summary of determinations of α_s using hadronic τ-decays (a), from lattice calculations (b), from DIS measurements (c) and from event shapes and jet production in e^+e^- annihilation (d) [7].

In DIS processes, the deviation from elastic scattering is described using the Bjorken scaling variable x. P is the four-momentum of the proton with the mass M. The quantities k and k' are the four-momentum of the incoming and outgoing lepton. The exchanged particle transfers the four momentum $q = k - k'$. The Lorentz invariant energy transfer is $\nu = \frac{Pq}{M_h}$.

$$x = \frac{Q^2}{2M_h\nu} = \frac{Q^2}{2Pq} \tag{2.16}$$

In early publications at an energy of around 3 GeV, it was shown, that the scaling behaviour x is independent of Q^2. If the scattering occurs between spin-$\frac{1}{2}$ particles, the measured structure functions F_1 and F_2 can be related through the Callan-Gross relation, which is only valid for fermionic objects.

$$2xF_1(x) = F_2(x) \tag{2.17}$$

The DIS experiments confirmed that quarks are fermions. At higher momentum transfers, the parton masses can be neglected and the Bjorken scaling variable x can be interpreted as the momentum fraction of the proton carried by one of the constituents. The momentum fraction carried by the parton i can be written as probability function $f_i(x_i)$. It was found that only half of the momentum of the proton is carried by the quarks. The question about the missing momentum has

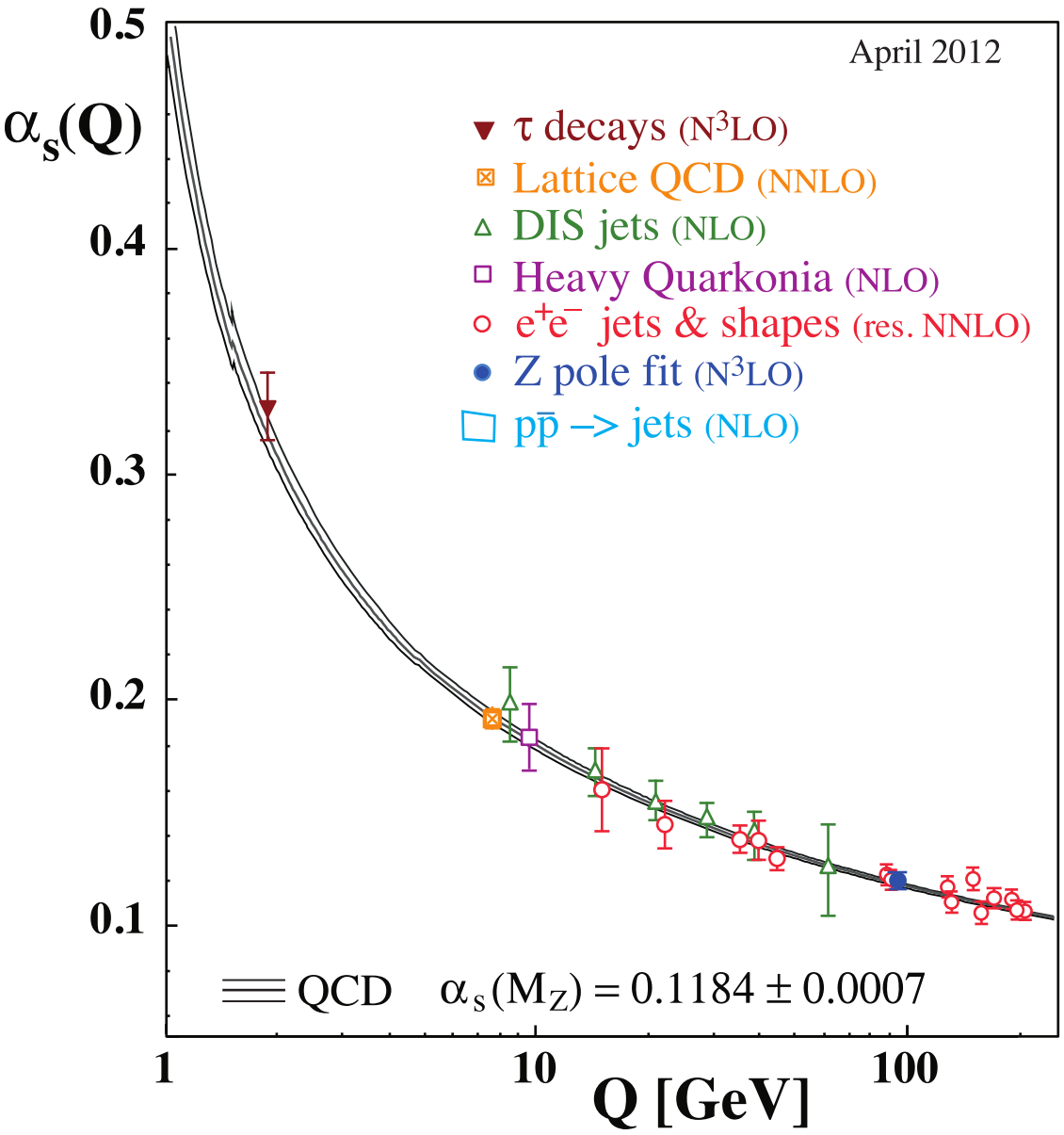


Figure 2.3.: Summary of measurements at different energies show agreement with the predicted running behaviour of α_s. The evolution to the Z boson mass scale yields the current world average value $\alpha_s(M_Z^2) = 0.1184 \pm 0.0007$ [7].

been solved by the QCD improved parton model which includes also the gluons as additional momentum-carrying partons.

2.2.4. QCD Improved Parton Model

In contrary to the naive assumptions of independent parton constituents in the quark parton model, the QCD improved parton model includes interactions between partons through the exchange of gluons. If the momentum transfer Q^2 is

large enough for α_s to be treated in a perturbative approach calculations can be performed.

However the cross sections cannot be completely calculated in a perturbative approach. An important property of the theory is the ability to factorise the cross section into a hard scattering matrix element and the parton distribution functions. With short-range hard matrix element $\hat{\sigma}$, which is calculable in perturbative QCD, and the long-range parton distribution functions f_a and f_b, the full cross section can be written as:

$$\sigma(x, Q^2) = \sum_f \int_{x_1,x_2}^1 dx_1 dx_2 f_a(x_1, \mu_F^2) f_b(x_2, \mu_F^2) \times \hat{\sigma}_{ab\to X}(Q^2, \mu_F, \mu_R) \tag{2.18}$$

However, through this separation, the PDFs is no longer scale independent, as gluons can be radiated. Dorkshitzer, Gribov, Lipatov, Altarelli and Parisi introduced the DGLAP evolution equations, which describe the impact on the PDFs due to changes in Q^2 [8, 9, 10]. The PDFs themselves cannot be calculated in perturbative calculations, but have to be derived from experimental measurements.

To compare the measurements at different scales, a PDF of a gluon or quark is defined at a low scale and then evolved to the experiment specific scale using the DGLAP evolution equations:

$$\frac{\partial q(x,\mu^2)}{\partial \log \mu^2} = \frac{\alpha_s}{2\pi} \int_x^1 \frac{dz}{z} \left(P_{q\to qg}(z, \alpha_s) q\left(\frac{x}{z}, \mu^2\right) + P_{g\to qq}(z, \alpha_s) g\left(\frac{x}{z}, \mu^2\right)\right) \tag{2.19}$$

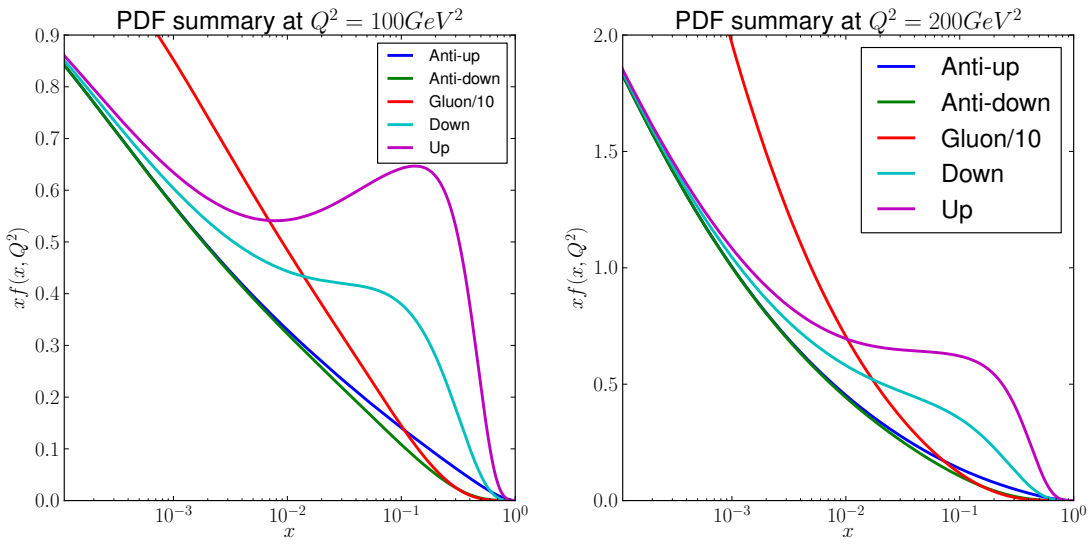
$$\frac{\partial g(x,\mu^2)}{\partial \log \mu^2} = \frac{\alpha_s}{2\pi} \int_x^1 \frac{dz}{z} \left(P_{q\to qg}(z, \alpha_s) q\left(\frac{x}{z}, \mu^2\right) + P_{q\to gg}(z, \alpha_s) g\left(\frac{x}{z}, \mu^2\right)\right) \tag{2.20}$$

z is the momentum fraction of the parton after radiation of the gluon, while the splitting functions $P_{a\to bc}(z, \alpha_s(\mu^2))$ describe the probability of radiation. The splitting function can be evolved in orders of α_s as

$$P_{ab}(x; \alpha_s(\mu^2)) = P_{ab}^0(x) + \frac{\alpha_s}{2\pi} P_{ab}^1(x) + O(\alpha_s^2) \tag{2.21}$$

The effect of the evolution can be seen from the comparison of a PDF at a low scale and a high scale. Figure 2.4 shows the fitted PDF at a scale of $Q^2 = 10\,\mathrm{GeV}^2$ and $Q^2 = 200\,\mathrm{GeV}^2$. The gluon splitting leads to a strong rise of the gluon PDF and also to a small increase of the quark PDFs in the low-x region when the PDF is evolved to higher scales.

As each experiment can only cover a specific region of the phase space, it is important to combine data from different experiments to measure the PDF with the highest possible precision. Currently the global PDF studies combine data of



(a) PDF at $Q^2 = 10\,\mathrm{GeV}^2$ (b) PDF at $Q^2 = 200\,\mathrm{GeV}^2$

Figure 2.4.: Parton distribution functions at different scales. Presented is an own fit produced with the HERAFitter framework using HERA DIS and CMS Inclusive jet data.

the HERA and the Tevatron colliders as well as data from fixed-targets experiments. The covered range in x and Q^2 of each experiment is shown in Figure 2.5.

Many different groups extract these PDF sets from published data using different approaches. The groups MSTW [12], CTEQ [13], NNPDF [14] and ABM [15] use data sets from all available experiments while the HERAPDF group [16] uses only data from the HERA collider. Due to the combination of datasets from many different sources, this poses also an important cross check of the compatibility between these data sets. More details on each PDF can be found in the cited publication by the respective PDF group.

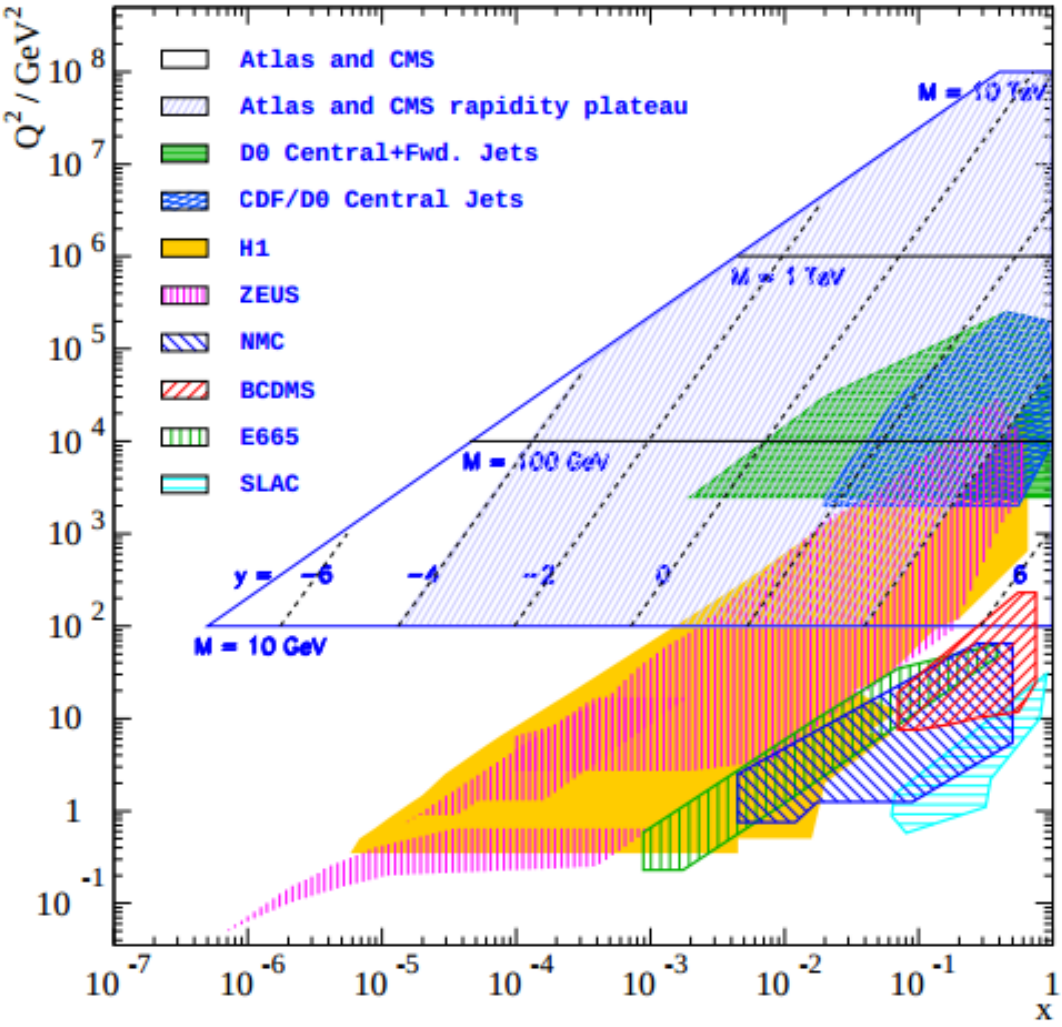


Figure 2.5.: Accessible phase space region for different experiments. The HERA data covers a wide range in x and Q^2 and therefore is included in all fits as base data set. LHC data allows to probe the PDFs at high scales and covers the remaining x region close to 1.0 [11].

2.3. Cross Sections

In order to verify the predictions made by pQCD, one needs an observable to compare the measurement to the calculation. The cross section σ is such an observable. It is defined as the interaction rate per target particle W normalised to the particle flux Φ:

$$\sigma = \frac{W}{\Phi} \tag{2.22}$$

The comparison between the theoretically derived and the measured cross section gives some indications of the goodness of the used model, or also might contain

hints on new physics. The commonly used unit of the cross section is a *barn*, with $1\,\mathrm{b} = 10^{-28}\,\mathrm{m}^2$. It is often useful to define the cross section differentially, e.g. in terms of transverse momentum p_T and the rapidity y as it has been done for the inclusive jet cross section.

The interaction rate W can be calculated with Fermi's Golden Rule, which uses the quantum-mechanical transition matrix elements $|M_{fi}|$ between initial and final state and the energy density ρ_f of the final states.

$$W = \frac{2}{\pi} \mid M_{fi}^2 \mid \rho_f \tag{2.23}$$

The properties of the collider are related to the number of produced events.

$$\sigma = \frac{W}{L} \tag{2.24}$$

Therefore, the instantaneous luminosity L is defined, which contains for the LHC quantities like the numbers of protons per bunch, the properties of the beam optics and the energy. The actually used luminosity is measured experimentally by relating the nuclear elastic forward scattering amplitude with the total cross section. The absolute luminosity scale is determined using Van der Meer scans. This procedure scans the LHC beams through one another to measure the size of the beams at the collision point.

2.3.1. Jet Cross Sections

In scattering processes with large momentum transfers, the outgoing partons manifest themselves as a collimated spray of particles in the detector. These structures are called *jets*. The appearance of jets is a fundamental feature of QCD and the study of the jet production allows precision studies of QCD predictions. The definition of a jet is crucial. The prescription for grouping particles together are called jet algorithms.

The properties of a jet algorithm are explained in the context of the anti-k_t algorithm, which is used in many recent LHC analyses. The anti-k_t algorithm uses the resolution parameter R for fixing the jet size. The distance measure $d_{i\mathrm{B}}$ and d_{ij} between two particles i and j in four-vector space

$$d_{ij} = \min(p_{\mathrm{T}i}^{-2}, p_{\mathrm{T}j}^{-2}) \frac{\Delta R_{ij}^2}{R^2} \tag{2.25}$$

$$d_{i\mathrm{B}} = p_{\mathrm{T}i}^{-2} \tag{2.26}$$

with the angular distance

$$\tag{2.27}$$

$$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 \tag{2.28}$$

$$\tag{2.29}$$

is used to decide which particles are clustered together. The pair-wise recombination algorithm starts with a list of input particles. At first the distances d_{ij} and $d_{i\mathrm{B}}$ are calculated. If the minimal distance value is given by d_{ij}, the particles i and j are combined into a new object which replaces the particle i and j in the input list. The algorithm restarts at the first step. If the smallest distance is $d_{i\mathrm{B}}$, the object i is removed from the input list and added to the list of final state jets. This procedure is repeated unless no more objects are in the input list.

The anti-$\mathrm{k_t}$ algorithm has the advantage that the resulting jet areas are fairly circular for harder jets, while soft jets will be crescent-shaped if they are located next to a hard jet. An important criteria for a good jet-algorithm is the collinear and infra-red safety. It must be ensured that an additional radiation of an infra-red soft gluon does not lead to the merging of two jets into one as shown in Figure 2.6a. Additionally, collinear safety must be guaranteed. If the transverse energy is split into two parts, e. g. by radiation or the granularity of the detector, the final observable must not change. If the transverse momentum is split, this could lead to the rejection of a jet as it is below a certain threshold. This is shown exemplary in Figure 2.6b.

An elementary and well-known check of QCD predictions can be done by the analysis of the inclusive jet production. The energy of the jet corresponds to the energy of the initial parton. This means that the jet spectrum yields information about the parton distribution within the proton. Most often the inclusive jet cross section is measured double differentially in terms of the absolute rapidity $|y|$ and the transverse momentum p_{T} of the jet. Theoretical calculations of the inclusive jet cross section are available in next-to-leading order (NLO) and the comparison with the measured cross section allows to put constraints on the PDFs as well as the determination of the strong coupling α_{s}. Since NLO calculations on parton-level cannot be compared directly to the unfolded data cross sections, additional non-perturbative corrections have to be applied which take hadronisation and underlying event effects into account.

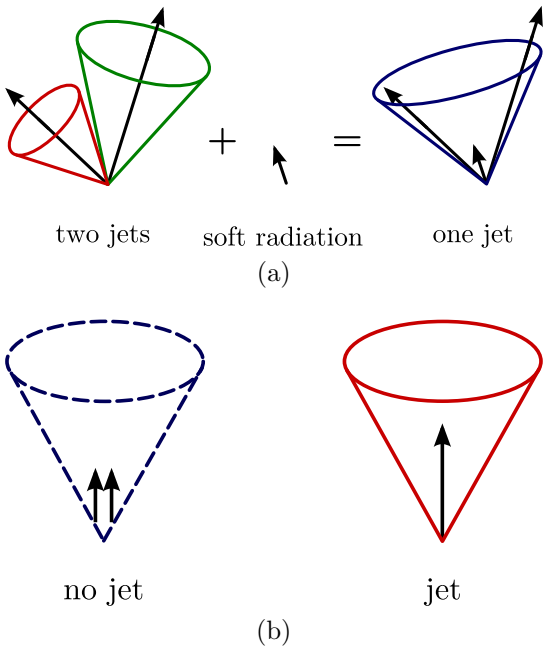


Figure 2.6.: A Jet algorithm must ensure that the additional radiation of a soft gluon does not result in the merging of two jets into one. This feature is called *infrared safety*. Additionally the algorithm must make sure that an additional collinear radiation does not change the output of the algorithm. This is referred to as *Collinear safety* [17].

3. The CMS Experiment

All particles predicted by the standard model have been found. Only the Higgs boson has veiled itself very well. Since the experiments performed at the LEP[1] and Tevatron did not show any evidence for the Higgs boson, a hadron-collider has been designed to investigate the remaining phase-space regions. The LHC would be either able to find the Higgs boson or to fully exclude a standard model Higgs boson below 1 TeV. A very detailed description of the LHC can be found in [18].

3.1. The Large Hadron Collider

As a consequence of the requirements to a new general purpose collider, the Large Hadron Collider (LHC) with a design centre-of-mass energy of 14 TeV at proton-proton collisions was the finally chosen model. The main goals of the LHC are the discovery of the Higgs boson and the search of physics beyond the standard model. The confirmation of super-symmetric theories or the production of dark matter candidates would help to solve problems not covered by the Standard Model. The wide range of never studied phase space allows precision studies to probe the standard model at higher energies than ever before. In November 2009, the LHC started to produce the first proton-proton-collisions. In the years 2010 and 2011, proton-proton collisions at a centre-of-mass energy of 7 TeV and in the year 2012 at 8 TeV have been recorded.

The LHC has been built inside the LEP tunnel at CERN[2]. It consists of about 5000 superconducting coils, which hold the beams on track in a tunnel of 27 km circumference. The accelerator consists of two rings in which the oppositely directed proton beams circulate. Before the injection and acceleration of protons in the LHC, they are pre-accelerated in the Proton-Synchrotron (PS) and the Super-Proton-Synchrotron (SPS) up to an energy of 450 GeV. Several more pre-accelerators are used in the full acceleration process which can be seen in Figure 3.1.

In a final step, 16 high frequency niobium cavities accelerate the beams to the desired energy. Along the ring, 1232 dipole magnets which generate a magnetic field of up to 8.3 T, bend the proton beams for the circulation inside the tunnel. There are also additional quadrupole and sextupole magnets which can focus and squeeze the beam.

[1] Large Electron-Positron Collider
[2] Conseil Européen pour la Recherche Nucléaire

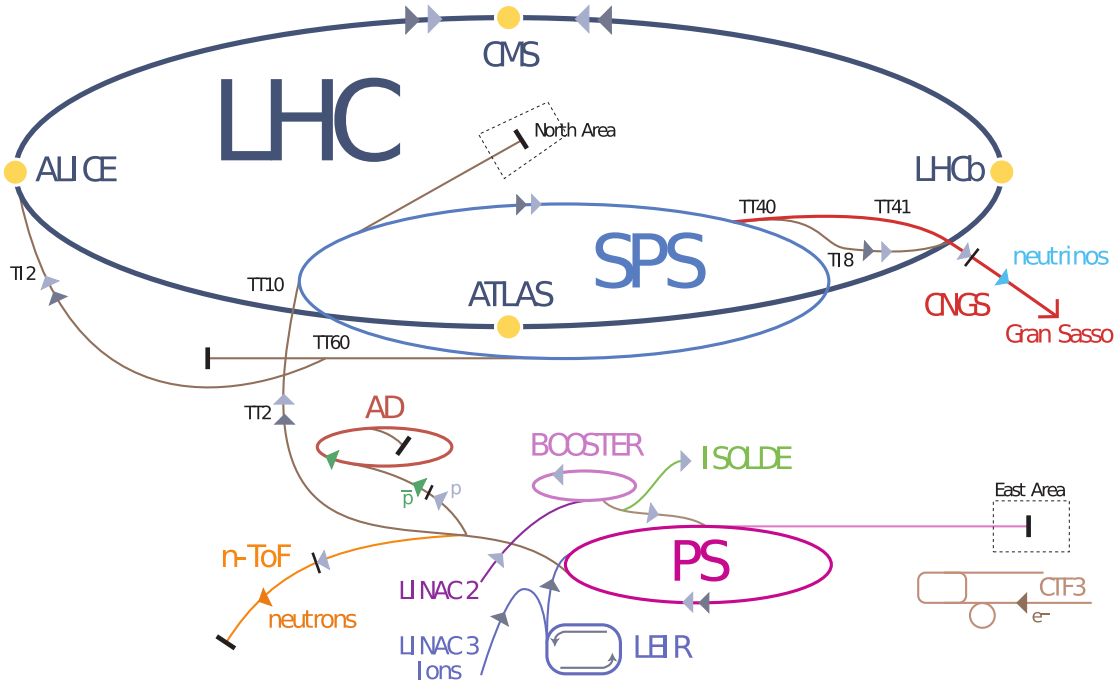


Figure 3.1.: Schematic view of the CERN accelerator complex. It shows the LHC ring with the four beam crossing points as well as the various pre-accelerators [19].

The two proton beams are brought to collision at four interaction points in the tunnel. At each collision point, one of the four big LHC experiments is installed: ALICE[20], ATLAS[21], CMS and LHCb[22]. The main goal of the ALICE experiment is the detailed study of quark-gluon plasma while LHCb is specifically designed for precision measurements of the CP violation and the decays of B mesons. Near the CMS experiment, the TOTEM[23] detector is installed. It is designed to study diffractive physics and helps to measure the total cross section. ATLAS and CMS are general purpose detectors with the aims of the discovery of the Higgs boson, the search for physics beyond the standard model and the confirmation of predictions by the standard model at high energies.

3.2. The Compact Muon Solenoid Experiment

The Compact Muon Solenoid (CMS) is one of the two general-purpose experiments at the LHC. The collaboration, which built and runs the detector, consists of approximately 3600 scientists, engineers and students from 38 countries.

The detector spans itself over a length of 21.5 m and a diameter of 15 m. It has a total weight of more than 12 000 t. The silicon tracking system and both the hadronic and the electromagnetic calorimeter are placed within a 13 m long superconducting solenoid, which provides a magnetic field of about 3.8 T. The

muon chambers are embedded in the iron return yoke outside the solenoid.

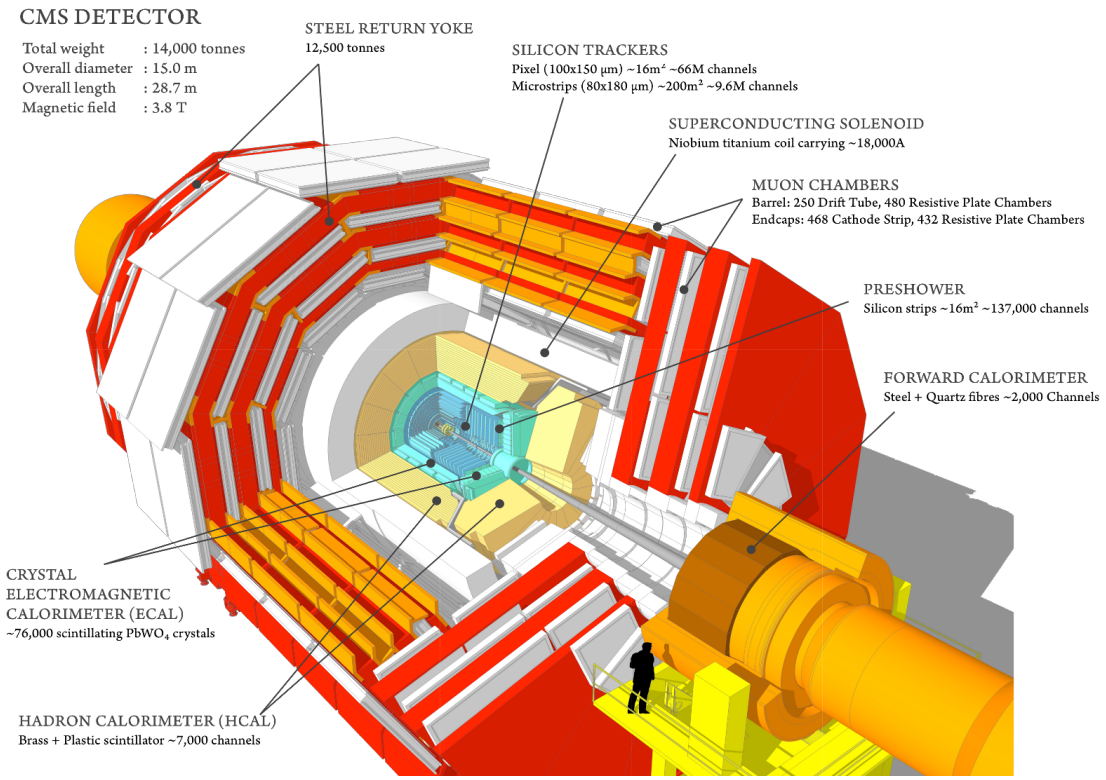


Figure 3.2.: Descriptive view of the CMS detector with the silicon detector system, the calorimeters within the solenoid and the muon detectors embedded in the return yoke [24].

3.2.1. Coordinate System

The coordinate system convention used by CMS has its origin at the nominal collision point. The x-axis points to the LHC ring centre and the y-axis points vertically upwards. The z-axis is aligned along the beam pipe, so that the coordinate system is right-handed.

Due to the geometrical form of the detector, it is beneficial to use cylindrical coordinates. r describes the distance of the point to the coordinate system origin, while ϕ is the azimuthal angle in the x-y-plane and θ the polar angle between the point and the beam pipe.

A more descriptive variable than the polar angle θ is the pseudo-rapidity η. The pseudo-rapidity is 0 for $\theta = \frac{\pi}{2}$ and increases while the angle to the beam direction gets smaller. The sign of η indicates the hemisphere. It does not depend on any knowledge about the mass of the particle.

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right) \tag{3.1}$$

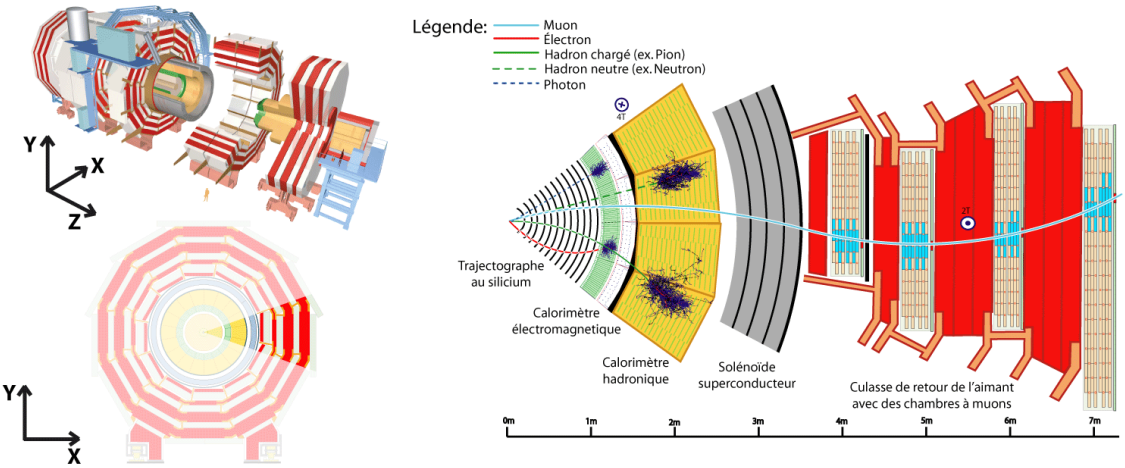


Figure 3.3.: A profile of the central detector with its various sub-detectors. Additionally the tracks of several particles and their interactions are shown. All particles but the muons and the neutrinos are stopped within the detector[25].

$$y = \frac{1}{2} \ln \frac{E - p_z}{E + p_z} \tag{3.2}$$

Another important quantity is the rapidity y. It is additive for Lorentz transformations along the beam axis. The rapidity and the pseudo-rapidity converge for very high energies $E \to |p|$ or for massless particles.

3.2.2. Silicon Tracking System

The silicon tracker is the sub-detector closest to the collision point. It is able to measure the trajectories of charged particles with a high precision and efficiency. It has a length of 5.8 m and a diameter of 2.5 m. Due to the high number of simultaneous collisions in the same bunch of protons, called pile-up events, a very fast response combined with a high granularity is necessary for the distinction of different collisions. Furthermore it is important to carefully choose the detector material to minimise the radiation damage and ensure a long lifetime. A schematic view of the inner tracking system can be found in Figure 3.4.

The beam line is surrounded by three cylindrical layers of pixel detectors. They are placed in the barrel region at the radii 4.4, 7.3 and 10.2 cm. The endcap region is covered by two disks on each side at a distance of 34.5 cm and 46.5 cm to the interaction point.

The huge number of 66 million detector pixel cells with a size of around 100 × 150 µm^2 deliver a spatial resolution of 15 to 20 µm. If charged particles pass the silicon pixel detector, they induce electron-hole pairs and a current can be measured.

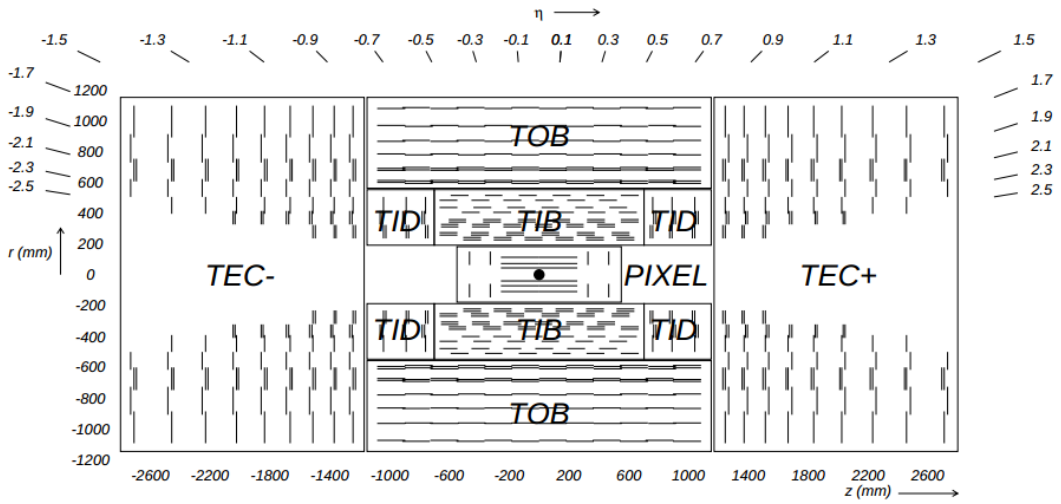


Figure 3.4.: The inner silicon tracking system in the r-z plane. The Tracker Inner Barrel (TIB),the Tracker inner Disks (TID), the Tracker Outer Barrel (TOB) and the Tracker EndCaps (TEC) are shown [26].

To ensure a good reconstruction of the track of the particle, the design is chosen in a way that at least three tracking points for each charged particle exist.

Since the particle density decreases with increasing radii, it is possible to replace the silicon pixel detectors by silicon strip detectors. The pixel detector is surrounded by ten strip detectors. These can be divided into three subsystems. The Tracker Inner Barrel (TIB) and the Tracker Inner Disks (TID), the Tracker Outer Barrel (TOB) and the Tracker EndCaps (TEC). The TIB consists of four layers located at 20 cm to 25 cm radii from the beam line while the TOB with six layers reaches up to a radius of 120 cm.

The 15400 modules with over 9.6 million silicon strips cover an area of 210 m^2. They typically measure more than 9 tracking points of a passing particle with a resolution 200 µm to 500 µm in the z- and 20 µm to 50 µm in the r-ϕ-region.

3.2.3. Electromagnetic Calorimeter

The calorimeters are designed to stop the particles so that the deposited energy can be measured. The electromagnetic calorimeter (ECAL) measures the energy of photons, electrons and to some extent also the deposited energy of passing charged hadrons, since they lose energy in a cascade of bremsstrahlung and electron-positron pair production processes. It has been specifically designed to deliver a very good diphoton energy resolution to measure the decay of a Higgs boson into two photons with high precision. A profile of the ECAL is shown in Figure 3.5.

When a high-energy particle collides with the nuclei of the ECAL crystals, a cascading shower of electrons, positrons and photons is generated. The atoms in

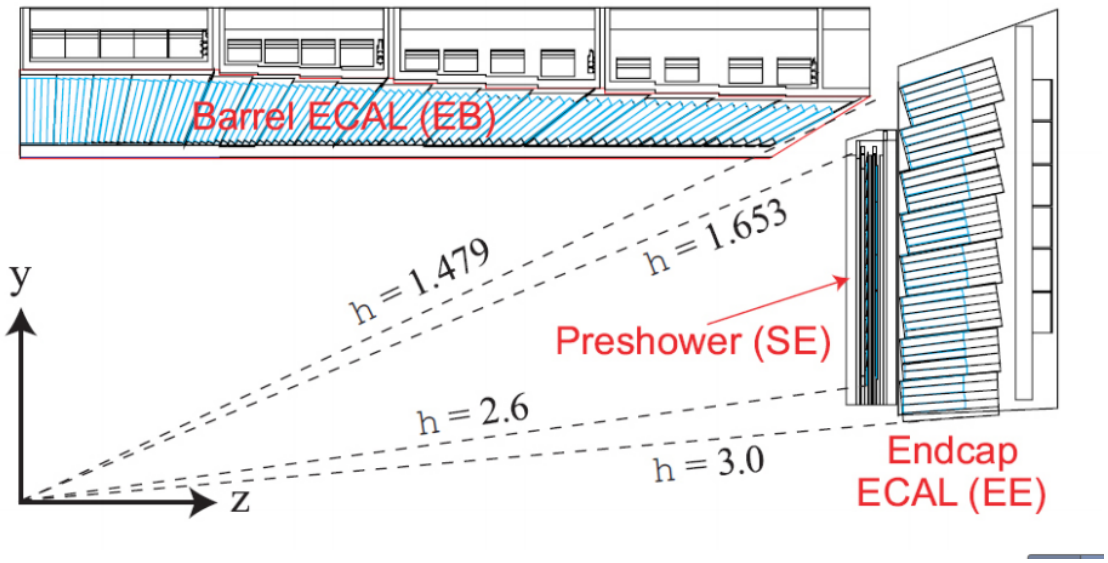


Figure 3.5.: Transverse section of the ECAL with the components. The ECAL is divided into the barrel (EB) and the endcap (EE) region. Additional pre-shower detectors (SE) improve the detector performance Figure from [27].

the crystals absorb some of the energy and the electrons are excited. When they relax again, they radiate scintillation light. The signal is amplified by additional Avalanche Photodiodes (APDs) which measure the signal in a photodetector after amplifying the current of the incident photon by as much as 100 million times. The measured number of photons is directly proportional to the energy of the passing particle. Due to the high hadron fluence in the endcaps, vacuum phototriodes (VPTs) are used.

In the CMS detector, lead tungstate crystals ($PbWO_4$) are used as both absorbing and scintillating material. This allows to build a compact detector without much blind material and a high granularity. The crystals combine a high density of $\rho = 8.28\,\mathrm{g/cm^3}$ with a short radiation length $X_0 = 0.89\,\mathrm{cm}$ and a Moliere radius of 2.2 cm. Furthermore lead tungstate has a very short scintillation decay time which helps to separate consecutive events. Additional pre-shower detectors (SE) consisting of lead as absorbing material and silicon detectors help to separate high energetic single photons from photons produced in pair production processes during τ^0 decays.

The ECAL is split into the barrel (EB) and the endcap (EE) region. The barrel region has a detecting volume of $8.14\,\mathrm{m^3}$ and weights more than 67 t. More than 60000 crystals allow a coverage of $|\eta| < 1.479$. The length of each crystal of 230 mm is equivalent to 25.8 radiation lengths.

The endcap sub-detector further extends the pseudo-rapidity coverage to $|\eta| < 3.0$. About 14500 crystals are mounted in a rectangular grid. The crystal length of 220 mm corresponds to 24.7 radiation lengths.

3.2.4. Hadronic Calorimeter

The goal of the hadronic calorimeter (HCAL) is to provide precise information about the deposited energy of hadronic particles, which are important for the study of hadronic jets and the missing transverse energy. Since the hadronic particles have a longer radiation length than photons or electrons, they are not stopped in the ECAL. The function of the ECAL applies in principle to the hadronic calorimeter (HCAL) as well. Unlike the ECAL, the HCAL is a sampling calorimeter. This means that the calorimeter consists of alternating layers of absorbing and scintillating material. As a result, only a part of the deposited energy is measured and the resulting energy has to be corrected.

The HCAL is divided into four sub-detectors. Two of the parts, the HCAL Barrel (HB) and the HCAL Endcap (HE), are located inside the solenoid and around the ECAL. As the size of solenoid limits the size of the HCAL and thus also the absorbing capabilities, the additional outer calorimeter (HO) was installed outside of the solenoid to measure the tails of high energetic jets. The fourth component is the forward calorimeter (HF), which extends the pseudo-rapidity coverage of the HCAL. The HB, HO and HE allow a detection of particles up to a pseudo-rapidity of $|\eta| < 3.0$ while the HF extends this up to $|\eta| < 5.0$. They are located at $z = \pm 11.2\,\mathrm{m}$. A schematic view of the hadronic calorimeter is shown in Figure 3.6.

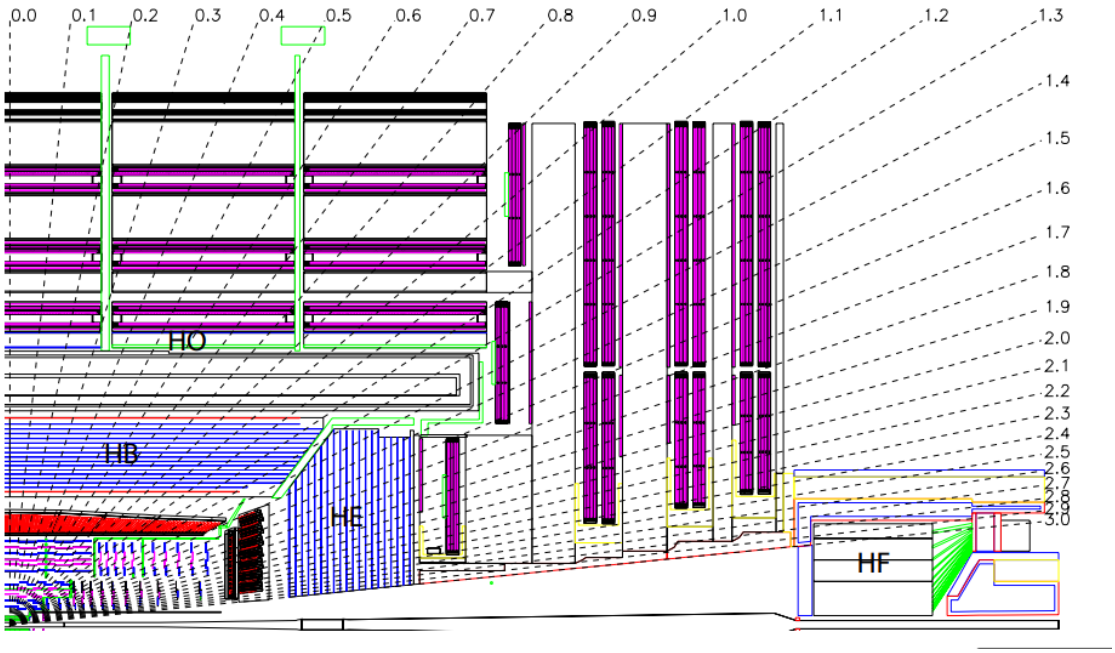


Figure 3.6.: The hadronic calorimeter system. The different sub-systems are the HCAL Barrel (HB), the HCAL Endcap (HE), the outer calorimeter (HO) and the very-forward calorimeters (HF) [26].

The HB and HE consists of brass as absorbing material which is arranged alternating to plastic scintillators. Brass has a density of $\rho = 8.53\,\mathrm{g/cm^3}$ and an interaction length of $\lambda_I = 16.42\,\mathrm{cm}$. Since the HF is not that much affected by the strong magnetic field, iron is used as absorber here. The absorbing capabilities vary for the different detector regions between 5.82 λ_I and 10.6 λ_I. Since this is not sufficient for a full screening of the high energetic hadronic showers, the solenoid coil acts as absorbing material for the HO. This leads to a total absorber thickness of 11.8 λ_I.

3.2.5. Superconducting Solenoid

The solenoid magnet used by CMS provides a longitudinal magnet field of about 3.8 T. The magnetic field is needed to bend the tracks of charged particles for the momentum measurement. As the inner detector components are very compact, they are placed within the magnetic coil. This reduces the jet energy uncertainty due to less dead material inside the calorimeters. The magnetic flux is guided by the iron return yoke, which also contains the muon chambers.

With a radius of above 3 m and a stored energy of 2.5 GJ, this solenoid is the largest superconducting magnet in the world.

3.2.6. Muon Chambers

The muon and neutrinos are the only particles which are able to leave the inner detector region. While neutrinos cannot be detected at all with the CMS detector and muons cannot be stopped in the detector, the muon momenta can be measured through ionisation in the inner silicon tracker and the outer muon chambers. The muon chambers are placed outside the superconducting solenoid inside the iron yoke.

In many analyses, like studies of the decay of the Higgs boson or searches for particles beyond the standard model, a precise knowledge on the involved muon is mandatory. CMS can identify and measure the muon momenta with a high precision due to a combination of the information from the silicon tracking systems and the muon chambers. A slide of the CMS detector with a detailed view of the muon system is provided in Figure 3.7.

The muon system consists of two parts, the barrel (MB) and endcap region (ME). The barrel part covers pseudo-rapidities up to $|\eta| < 1.2$ while the forward part extends this up to $|\eta| < 2.4$. The muons are detected using gaseous detectors. In the barrel region, aluminium drift tube chambers (DT) are used. Although they have a slow response time they were preferred due to their good spatial resolution. This is possible because of the low muon rates in the central region.

In contrast to the barrel region, the forward regions are challenged by high muon rates and a strong magnetic field. Therefore, drift tube chambers are not suitable.

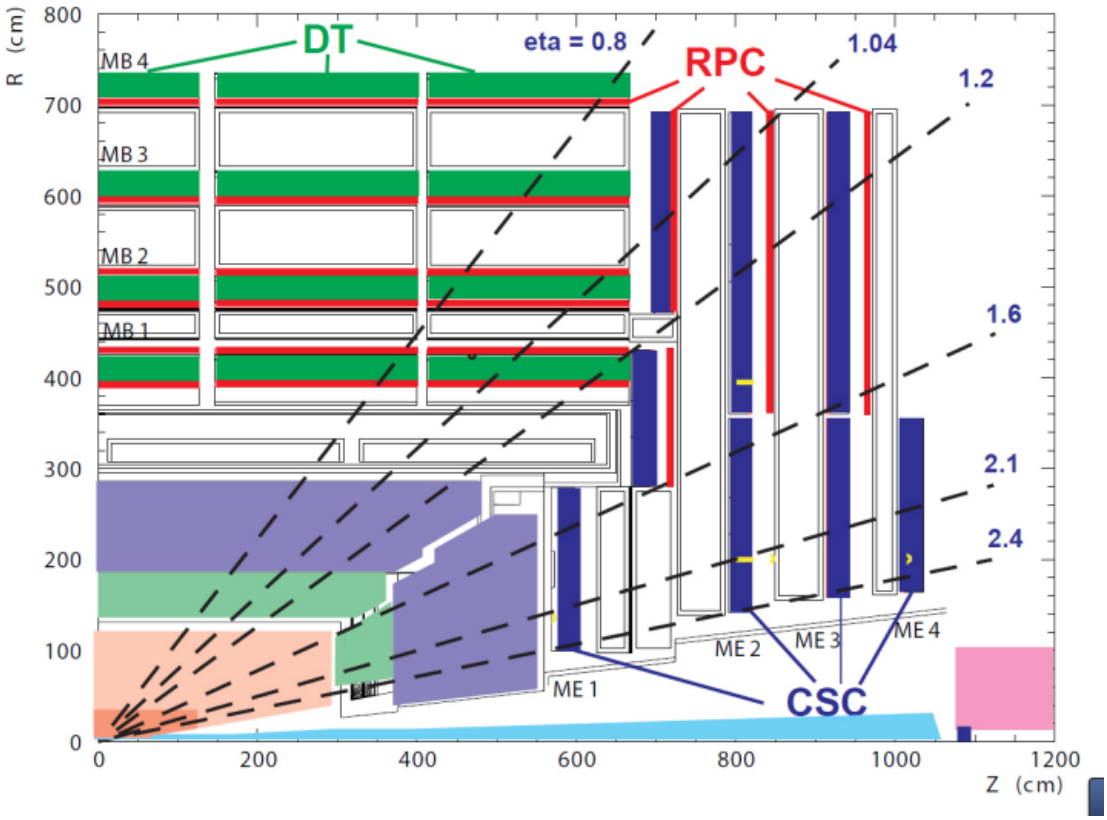


Figure 3.7.: Section of the CMS detector with respect to the muon system. The muon system is divided into the the barrel (MB) and the endcap (ME) region. Several used detectors are the drift-tube chambers (DT), the cathode strip chambers (CSC) and the resistive plate chambers (RPC) [27].

Cathode strip chambers (CSC) are used in the endcaps since they combine a good spatial resolution with a good time resolution. In both the central and the forward region, additional resistive plate chambers (RPC) are used. This sub-detector has a very good time resolution and a fast response time at the expense of a slightly worse spatial resolution. Due to their fast response, their information is used for the Level-1 trigger.

3.2.7. Data Acquisition and Trigger

There are about 10^9 proton-proton collisions every second at the design luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. As each collision event has an average size of 1 MB, this huge amount of data cannot be stored permanently with current storage technologies. Many of the collision events are the product of low-energetic processes which are of less interest to most analyses. Therefore several trigger levels are applied to extract potentially interesting events and hereby reduce the total amount of data.

The Level-1 trigger is implemented in hardware. With information of the calorimeters and the muon system, the Level-1 trigger reduces the event rate from 40 MHz to 100 kHz. After this hardware trigger, the events have to pass the so called High Level Trigger (HLT), which is a software trigger able to use information of all detector components for its decision. This trigger further reduces the rate from 100 kHz to about 600 events per second, which can be handled by the storage system.

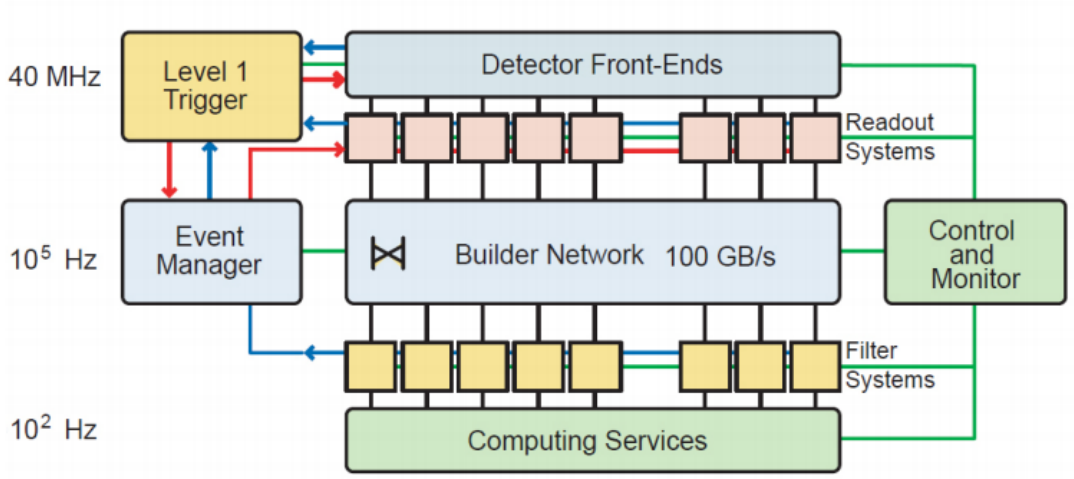


Figure 3.8.: Architecture of the CMS trigger and data acquisition system. It consists of readout electronics, a L1 hardware trigger and an additional software based trigger (HLT). It reduces the event rate of 40 MHz to about 600 Hz [26].

The storage and the analysis of these huge amounts of data is one of the challenges in the era of the LHC. The Worldwide LHC Computing Grid (WLCG) tries to satisfy the requirements by providing a sophisticated distributed computing infrastructure. The WLCG structure from the CMS view is shown in Figure 3.9.

There is one Tier-0 located at the CERN facility which stores the raw data from the detectors and provides prompt reconstructed samples. The data is then distributed to the different Tier-1 sites, where they can be reprocessed and skimmed with huge computing farms. The connected Tier-2 sites allow the analysis groups to perform their analyses using the data. They also provide the computing infrastructure for the simulation of collision events which is necessary for comparisons between experimental measurements and theoretical calculations.

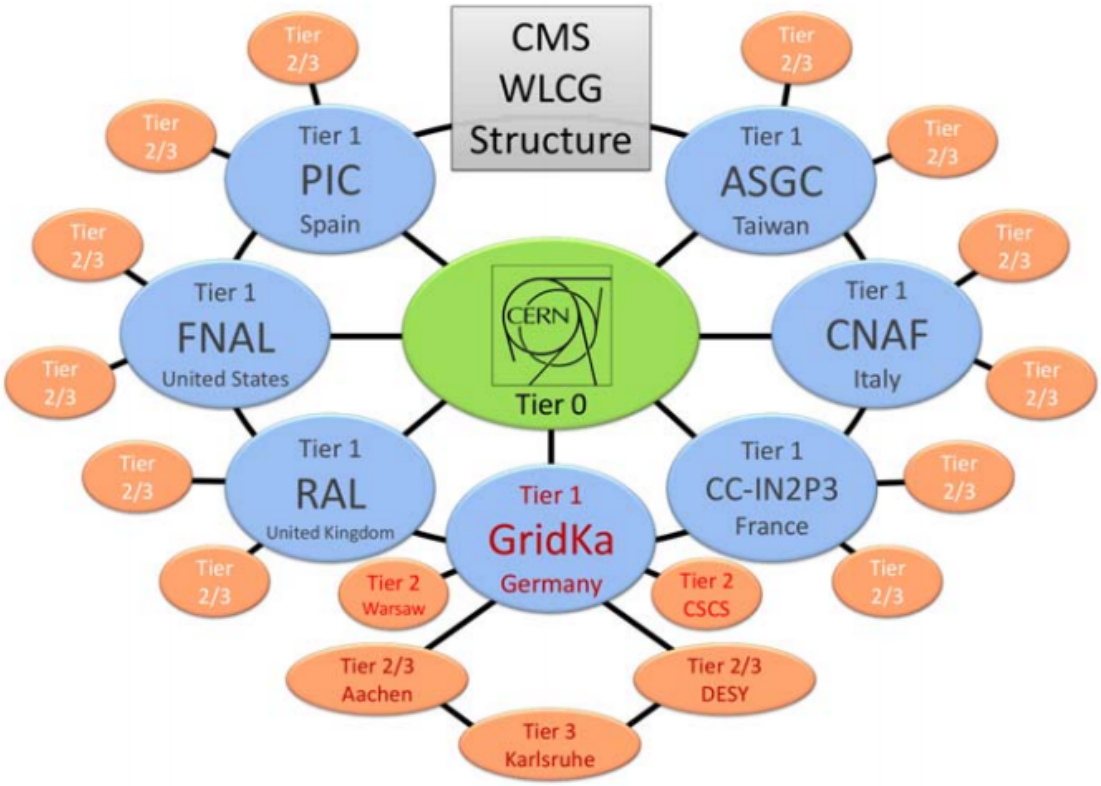


Figure 3.9.: The WLCG structure from CMS point of view. The attached Tier-2 and Tier-3 in Germany are shown as wells [28].

4. Analysis Tools

Today's experiments in particle physics have become very complex and produce enormous amounts of data. This poses challenging requirements to the software tools and work-flows on the computing side. Multiple tools are needed for the simulation and the processing of generated or recorded events. Theorists provide software packages capable of simulating cross sections for many processes to compare the models with experimental measurements. To handle the huge amount of data and needed computing capacity, the computing tasks are divided in separate jobs and are submitted to local batch resources or the Grid. This approach allows to parallelise the processing performed by these jobs to a high degree.

4.1. Standalone Tools

In this section, frequently used software tools for the analysis will be explained. They are used for the processing of the data as well as the visualisation.

4.1.1. ROOT

ROOT is an object-oriented analysis software framework which has been designed to efficiently analyse the huge amount of data expected at the LHC. ROOT has been developed at CERN and is executable on all major software platforms. It is implemented in C++ and ROOT programs can be either compiled natively or use the included C++ interpreter CINT.

While ROOT was initially designed to store a huge amount of data in an object-like structure and access them without reading the bulk of data, it also includes several classes for common tasks like curve fitting, histogramming, multivariate data analysis and visualisation [29]. A detailed description of the ROOT software framework can be found in [30].

4.1.2. matplotlib and NumPy

Most of the analysis steps in this study were performed with the scientific computing package NumPy within the scripting language Python. NumPy supports users with efficiently processing multi-dimensional data. It was combined with matplotlib for the visualisation of most of the plots in this thesis. matplotlib is a 2d plotting library for Python which produces publication quality plots in a variety

of formats. It offers an easy to use object-oriented interface with extensive options for customisation.

4.1.3. grid-control

All necessary data processing steps and fits were submitted to a batch system. The job submission tool grid-control is able to submit jobs to both local and grid computing resources in an efficient way [31]. grid-control also supports parallelised jobs, which are necessary for multidimensional fits comprising many input parameters.

4.2. HERAFitter Framework

The HERAFitter framework is an initiative from the H1 and ZEUS collaborations to provide an open source fitting framework combining tools from theorists and experimental physicists. HERAFitter was designed to derive the parton distribution functions of the proton from data of the HERA collider. The resulting PDF fits are also made publicly available [32, 33, 34].

However, HERAFitter is not limited to processing data from HERA, but it empowers the user to include own datasets from a variety of experiments including fixed-target experiments and data from hadron colliders.

Recent analyses from the LHC perform measurements which further constrain the PDFs of the proton. Within the HERAFitter framework it is possible to test these constraints by including the new data in a combined fit. This is especially valuable because the compatibility of several datasets from different colliders can be studied in a simultaneous fit.

It is currently employed by the ATLAS and CMS collaborations to study the influence of jet measurements and W-asymmetry measurements on the parton distribution functions. These measurements help to understand the flavour composition of the proton and the role of the gluon in particular.

4.2.1. Design

HERAFitter is based on a modular design capable to include new modules easily. Theorists provide tools for the calculation of the cross sections of the processes. These tools are exploited in the HERAFitter framework for theoretical predictions and comparisons with the data as shown in Figure 4.1. The most important theory tools are:

QCDNUM provides the calculations of the DIS structure functions. The structure functions are calculated as a convolution of the parton distribution functions with the heavy-quark coefficient functions. Furthermore QCDNUM is employed for the evolution of the PDFs

fastNLO and APPLGRID provide access to pre-calculated tables with the perturbative coefficients. The perturbative coefficients can be convoluted with the PDFs in milliseconds to give the jet cross sections. Otherwise jet measurements could not be included in PDF fits, since a full calculation at sufficient accuracy is very time consuming.

Hathor is capable of calculating the $t\bar{t}$ cross section up to approximate NNLO accuracy.

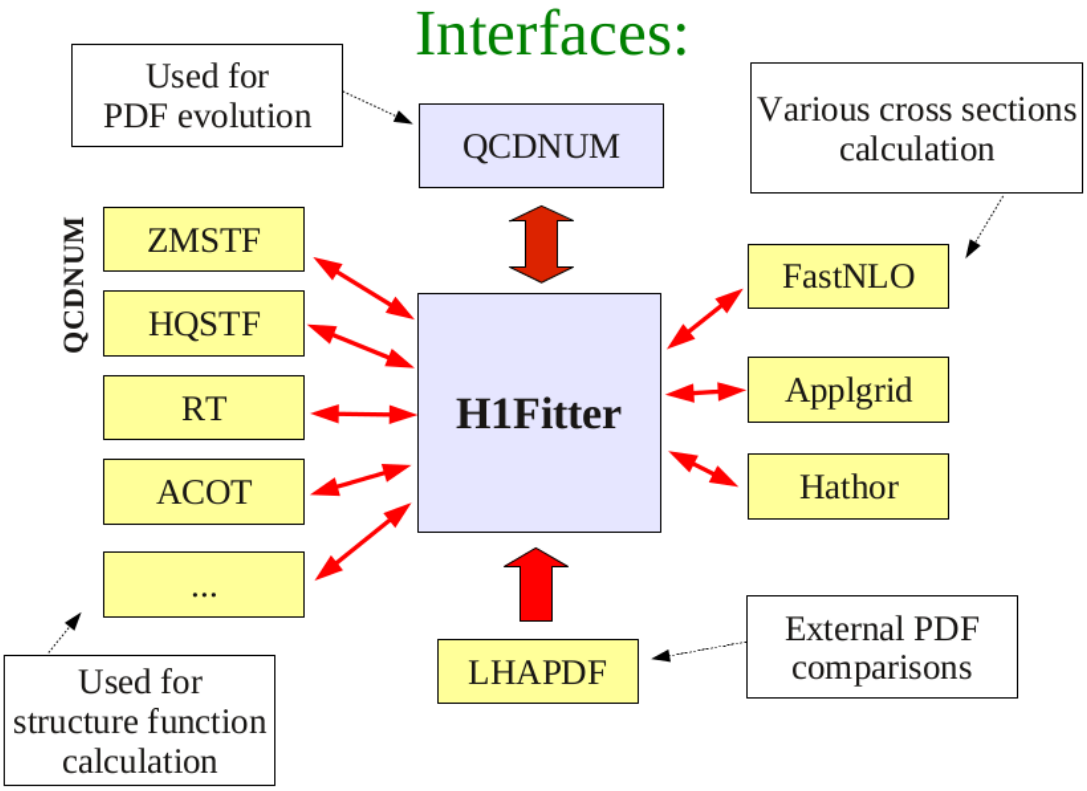


Figure 4.1.: Structure of the HERAFitter framework, here still referred as H1Fitter, showing the various theory tools. QCDNUM is used for the evolution of the PDFs and α_s. fastNLO, APPLGRID and Hathor are cross section integrators for jet and $t\bar{t}$ processes. LHAPDF is a PDF library, which is capable of interfacing all PDFs from the major PDF fitting groups [35].

HERAFitter provides several possible styles for the parametrisation of the PDFs at the starting scale of the fit.

The HERAPDF style parametrises the valence distributions, the gluon distribution and the u-type and d-type anti-quark distributions with a functional form

$$xf(x) = Ax^B(1-x)^C(1+Dx+Ex^2) \tag{4.1}$$

The CTEQ style is more flexible by extending the standard functional form $xf(x) = x^B(1-x)^C$ with additional terms.

$$xf(x) = a_0 x^{(a_1+n)}(1-x)^{a_2} e^{a_3 x}(1 + e^{a_4 x} + e^{a_5 x^2}) \tag{4.2}$$

These parametrisation styles are implemented in HERAFitter and can be chosen according to the needs of the user.

The minimisation itself is performed using the Fortran package MINUIT which was developed at CERN. It is widely used in particle-physics and regarded as very stable. Although there is a newer MINUIT version rewritten in C++, HERAFitter relies on the Fortran version. The minimised function is an extended χ^2 taking into account systematic and statistical uncertainties and their correlations.

4.2.2. Fitting Procedure

This section gives a brief overview how HERAFitter performs a PDF Fit. The more detailed description as well as the determination of the HERAPDF set is given in Section 5.3. The procedure can be divided into the following five steps.

Parametrisation at starting scale The starting scale as well as the parametrisation of the PDFs is chosen arbitrarily. The starting scale is most often chosen to be below the mass threshold of the charm-quark. HERAFitter provides several possible parametrisation schemes while the standard one is an extended polynomial function.

Evolution of the PDF In order to be convoluted with the hard matrix elements in the cross section calculation, the scale-dependent PDFs first have to be evolved to the scale of the measurement.

Calculation of the Cross Sections The calculation of cross sections at next-to-leading order is often very time-consuming. Special programs like fastNLO or APPLGRID ease this step by providing pre-calculated tables of the perturbative coefficients.

Best-fit PDF The calculated cross section is compared to the unfolded measurement. The goodness is estimated using a χ^2 estimator. The PDF parameters are fit until a best-fit PDF is found.

Estimation of Uncertainties There are several sources of uncertainty which affect the PDFs. The estimation of the uncertainties is performed either with a Monte Carlo approach or with the Hessian Method. The Monte Carlo approach is based on multiple fits with pseudo-data, while the Hessian Method propagates the uncertainty of the parameters on the PDF using a special technique described in Section 5.4.1

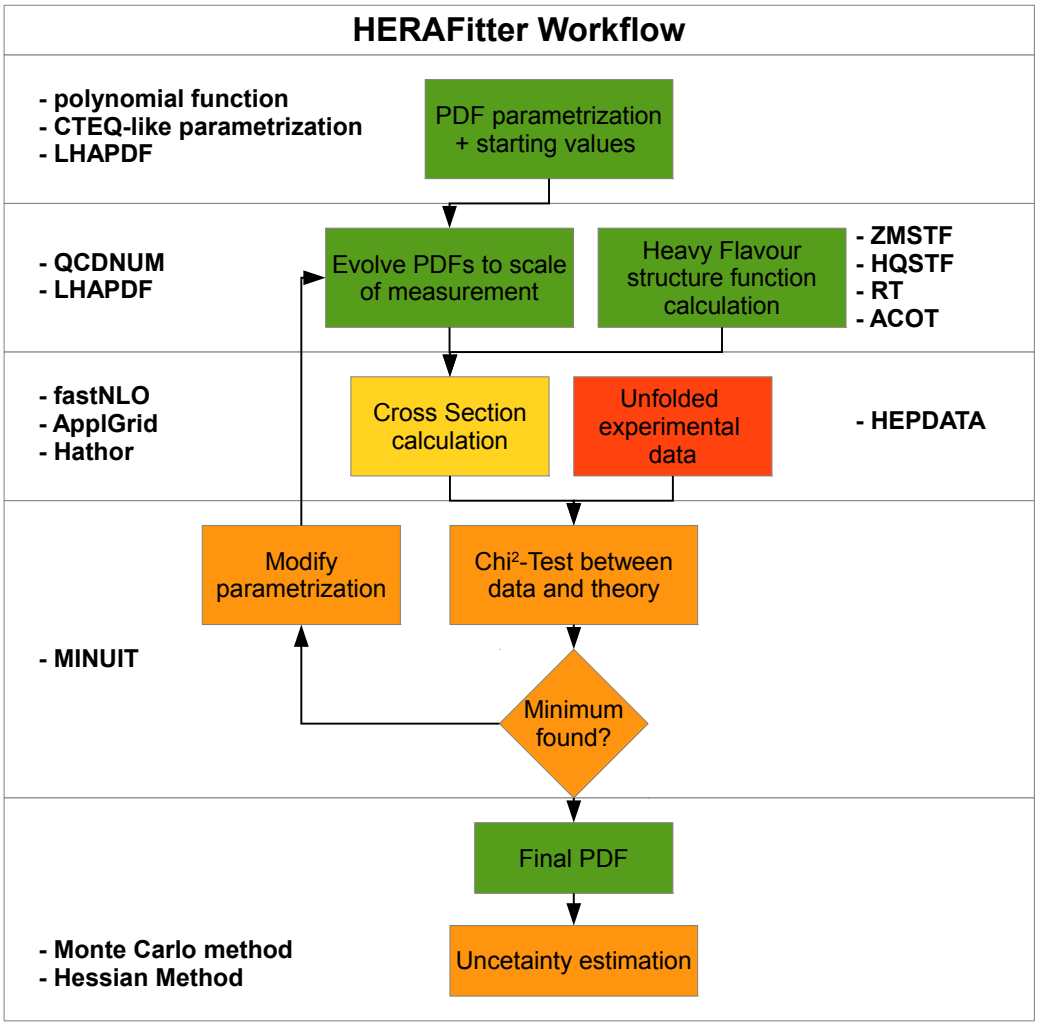


Figure 4.2.: The necessary steps for the determination of the PDFs. At first the parametrisation has to be chosen at a starting scale. This PDF is evolved to the scale of the measurement and the cross section of the process is calculated. The calculated cross section is compared to the unfolded data of the measurement and the parametrisation of the PDF is adapted to describe the data as accurately as possible. The best-fit PDF is the final PDF set.

The detailed workflow is shown in Figure 4.2. The settings of the HERAFitter software can be customised with Fortran steering cards. This allows users to adapt the software to their needs. It is also possible to interface fixed PDF sets using the LHAPDF [36] library for special studies. This is especially useful for the comparison of fixed PDF sets with self-made fits.

4.2.3. QCDNUM

QCDNUM is a highly efficient Fortran program for the QCD evolution of parton distribution functions. QCDNUM solves the evolution equations for parton distribution function and the strong coupling up to next-to-next-to-leading order in powers of α_s [37]. The integrated convolution mechanism allows the calculation of the deep inelastic structure functions. QCDNUM is interfaced to HERAFitter which provides the parton distribution functions.

4.3. NLOJET++ and fastNLO

NLOJET++ is a QCD cross section integrator, which can calculate cross sections for one-, two- and three-jet observables at next-to-leading order [38].

For the inclusion of the jet cross-sections in the PDF fits and for a measurement of the strong coupling, it is very important to calculate the cross section at NLO to reduce the scale uncertainties as much as possible. NLOJET++ uses the the Catani-Seymor dipole subtraction formalism to avoid infrared divergences, which can appear in these NLO calculations. The necessary computations are extremely time-consuming and not suited for PDF fits, in which the cross section has to be recalculated in each iteration.

The necessity of fast NLO calculations with slightly different PDFs or different values of α_s lead to the development of the fastNLO framework [39]. This program performs a separation of the parton distribution functions from the hard scattering matrix element by interpolating the PDF on a fixed grid in the fractional proton momentum space. The perturbative coefficients are calculated and then are stored in a table. This calculation with high statistics is challenging but has to be performed only once. Afterwards the perturbative factors can be folded with the PDFs and α_s. This procedure is very fast and can be performed in each iteration of a PDF fit. The fastNLOReader can use the evolution code provided by the LHAPDF library or can be interfaced with user code as shown with the QCDNUM package.

fastNLO is employed to calculate the predictions of the inclusive jet cross section and is used as well for the calculation of the PDF and scale uncertainties and the comparisons of different PDF sets.

5. Constraints on PDFs and Determination of α_s from Inclusive Jet Cross Sections

The parton distribution functions of the proton cannot be calculated from perturbative QCD, but have to be derived from comparisons of measured cross sections to theoretical predictions. The uncertainties on the PDFs is one of the dominant uncertainty sources in some analyses at high energies at hadron colliders. Therefore it is necessary to extract the PDFs with high precision by including many different data sets from various phase-space regions and processes. The LHC can reach kinematic regions never explored before and helps to further constrain the PDFs in these regions as was shown in Figure 2.5.

The aim of this study is to evaluate the impact of inclusive jet data from the CMS experiment on the PDFs. These PDF studies are performed using HERA-I DIS and CMS inclusive jet data in a combined fit. The second step is the extraction of the strong coupling from the inclusive jet data using PDF sets for a series of fixed values of α_s and corresponding NLO calculations.

5.1. Extraction of Parton Distribution Functions

The first step to determine the parton distribution functions is to define a parametrisation at a starting scale Q_0^2. The quark and gluon PDFs cannot be fitted separately as sum rules apply, e. g. the quark number or the momentum sum rule. The general approach is a parametrisation similar to the one chosen for the HERAPDF [16] determination.

$$xf(x) = Ax^B(1-x)^C(1+Dx+Ex^2) \tag{5.1}$$

The shape of the function is defined by a polynomial function in which the parameter B constrains the region $x \to 0$ and the parameter C the region $x \to 1$. The normalisation due to the quark number and momentum sum rule is done using the parameter A. Additional tuning is possible through the parameters D and E.

Since there is one PDF for the gluon and each quark and anti-quark, thirteen PDFs have to be determined. Heavy quarks are produced from gluons splitting

into heavy-quark pairs at energies above the mass threshold. Unlike the light-quark PDFs, these are not fitted to the data, but are generated in the process of the DGLAP evolution [40]. There exist different schemes for the calculation of the heavy quark contribution. As there are only a few experiments which can reach the scale of the top quark, this PDF is often neglected. Since the heavy quarks are not fitted to the data, seven PDFs remain. This number can be reduced further by assumptions on similarities between the two light down-type quarks d and s.

With a parametrisation at the starting scale, the DGLAP equation is used to evolve the PDF to the scale of the measurement. The next step is then to compare the corresponding cross section to the measurement using a least-squares estimator. The least-squares method is deployed to consider uncorrelated and correlated uncertainties. This step is repeated iteratively until a best-fit χ^2 is found while adapting the PDF parametrisation. The best-fit parametrisation defines the central PDF.

The final step is to estimate the uncertainty. This can be done using Monte-Carlo approaches as used by the NNPDF collaboration or methods like the Hessian method [41], in which a transformation to the eigenvector space is done. The Hessian approach employed here will be presented in detail in Section 5.4.

5.2. Parameter Estimation Based on Extended χ^2-Minimisation

The agreement between the data points D_i and the theoretical values T_i is estimated via a least-squares method. A χ^2 estimator for one experiment with N points of data including statistical and uncorrelated systematic sources of uncertainty σ can be written as

$$\chi^2 = \sum_i^N \frac{(D_i - T_i)^2}{\sigma_{\text{stat},i}^2 + \sigma_{\text{unc},i}^2}. \tag{5.2}$$

As this standard form does not consider any kind of correlations, most often the χ^2 is written in terms of the inverse covariance matrix V^-1 in which all bin-to-bin correlations can be considered:

$$\chi^2 = \sum_{ij}^N (D_i - T_i) V_{ij}^{-1} (D_j - T_j). \tag{5.3}$$

It is important to distinguish between additive and multiplicative uncertainties. If the relative multiplicative uncertainties are multiplicated with the data, lower values of the minimised quantity are preferred, due to the fact that the uncertainty stays the same. The application of the multiplicative uncertainties sources to the theory ensures that the uncertainties decrease if the theoretical prediction gets

smaller. This shift is the so-called d'Agostini bias [42]. The solution is to apply the multiplicative uncertainties to the theory. However the covariance matrix and its inverse has to be recalculated at each iteration of the fit in this approach.

The inversion of huge matrices is very time-consuming and can lead to numerical instabilities. It is often more convenient to combine the fully correlated systematic uncertainties analytically. This is performed using nuisance parameters r_k [43], which shift the prediction or the measurement. Since these nuisance parameters are determined separately, the influence of each systematic uncertainty source on the fit can be studied.

The χ^2 can be written as

$$\chi^2 = \sum_i^N \frac{(D_i - T_i(1 + \sum_k r_k \beta_{ik}))^2}{\sigma_{i,\text{stat}}^2 + \sigma_{i,\text{uncor}}^2} + \sum_k^K r_k^2. \tag{5.4}$$

taking into account the fully correlated sources of systematic uncertainty β_{ik} and the nuisance parameters r_k. These nuisance parameters must be determined to minimise the χ^2. This can be done by adding one free parameter for each nuisance parameter in the fit. This is not applicable because of the high number of sources of systematic uncertainty in the fitted datasets. Since the χ^2 is only a quadratic polynomial in r_k, the determination of the nuisance parameters can be done analytically. The minimum of Equation 5.4 can be found analytically by solving

$$A_{kk'} r_k = B_k \tag{5.5}$$

with the $K \times K$ coefficient matrix $A_{kk'}$ written as

$$A_{kk'} = \delta_{kk'} + \sum_i^N \frac{\beta_{ik}\beta_{ik'}}{\sigma_{i,\text{stat}}^2 + \sigma_{i,\text{uncor}}^2}. \tag{5.6}$$

and the vector B_k with K components for K sources of systematic uncertainty,

$$B_k = \sum_{i=1}^N \frac{\beta_{ik}(D_i - T_i)}{\sigma_{i,\text{stat}}^2 + \sigma_{i,\text{uncor}}^2}, \tag{5.7}$$

The nuisance parameters r_k are obtained by calculating the inverse of $A_{kk'}$ and multiplying with B_k

$$r_k = \sum_k^K A_{kk'}^{-1} B_k. \tag{5.8}$$

The shift of the nuisance parameter is taken into account in the equation by adding quadratically to the χ^2. This means that a shift of e. g. 2σ on the theory results in an increase of 4 of the χ^2. It was shown in [44] that the resulting Equation 5.4 is equivalent to the Equation 5.3 based on the full covariance matrices. This

equation is an appropriate definition of χ^2 for data with correlated systematic uncertainties.

5.3. Determination of the HERAPDF Set

The HERAPDF collaboration combined the measurements of neutral current (NC) and charged current (CC) e^+p and e^-p cross sections from the H1 and ZEUS experiments to reduce the systematic uncertainties by cross-calibration. The combined dataset was used as the sole input in the fit of the HERAPDF 1.0 set using the extended least-squares-fitting method as described in Section 5.2. A detailed description of the HERAPDF determination can be found in [32, 33, 34]. A derivation of the uncertainty on the PDF using the standard $\Delta\chi^2 = 1$ prescription resulted in a believable estimation of the uncertainty and demonstrated the consistency of the datasets. There are in total 110 systematic uncertainty sources which were treated using the Hessian procedure while the uncertainties of the combination procedure were treated using the Offset method [45]. The Offset method does not include the correlated systematic uncertainty sources in the fit but the data-points are offset to account for each systematic source and a new fit is performed for each of the variations. The deviations from the central fit are then added in quadrature.

This results in a χ^2 per degree of freedom of 574/582. The starting scale Q_0^2 at which the fit was performed was chosen to be $Q_0^2 = 1.9\,\mathrm{GeV}^2$ to be below the mass threshold of the charm quark. The heavy quark coefficient functions are calculated by the variable-flavour-number scheme RT [46, 47]. The input settings for the fit were chosen as follows

- As the input data from HERA have no direct constraint on the strong coupling constant, α_s is not included in the fit as a parameter but fixed to the value of $\alpha_s(M_z^2) = 0.1176$. This was the world average value of α_s at the time of the HERAPDF 1.0 determination.
- The heavy quark masses were chosen as $m_c = 1.4\,\mathrm{GeV}$ and $m_b = 4.75\,\mathrm{GeV}$. This is done to be consistent with the MSTW PDF analysis.
- A minimum Q^2 cut was set to $Q^2_{\mathrm{min}} = 3.5\,\mathrm{GeV}^2$. This is necessary to ensure the applicability of perturbative calculations in QCD.

The PDFs are parametrised at the starting scale using the following general parametrisation [48].

$$xf(x) = Ax^B(1-x)^C(1 + Dx + Ex^2) \tag{5.9}$$

A first fit is performed without the parameters D and E and they are only added if a significant influence on the χ^2 is observed. The HERAPDF fits determine

5 PDFs simultaneously, the valence-quark distributions xu_v and xd_v, the anti-quark distributions $x\bar{U}$ and $x\bar{D}$ and the gluon distribution xg. The anti-quark distributions are defined as $x\bar{U} = x\bar{u}$ and $x\bar{D} = x\bar{d} + x\bar{s}$. It was found that the best fit is performed using the following parametrisation with 10 free parameters.

$$xg(x) = A_g x^{B_g}(1-x)^{C_g} \tag{5.10}$$
$$xu_v(x) = A_{u_v} x^{B_{u_v}}(1-x)^{C_{u_v}}(1+E_{u_v}x^2) \tag{5.11}$$
$$xd_v(x) = A_{d_v} x^{B_{d_v}}(1-x)^{C_{d_v}} \tag{5.12}$$
$$x\bar{U}(x) = A_{\bar{U}} x^{B_{\bar{U}}}(1-x)^{C_{\bar{U}}} \tag{5.13}$$
$$x\bar{D}(x) = A_{\bar{D}} x^{B_{\bar{D}}}(1-x)^{C_{\bar{D}}} \tag{5.14}$$

The fit was performed with the following constraints: The normalisation parameters A_g, A_{u_v} and A_{d_v} are determined using the QCD quark number and momentum sum rules. Extra constraints on the low-x behaviour are given by fixing $B_{\bar{U}} = B_{\bar{D}}$ and $B_{u_v} = B_{d_v}$. This ensures that $x\bar{u} \to x\bar{d}$ for $x \to 0$. The strange quark PDF is not fitted, but determined as a fraction of $xf_{\bar{D}}$ as $x\bar{s} = f_s x\bar{D}$. The value chosen for f_s is 0.31. This value results from the determined fraction using neutrino-induced di-muon production. The result of this fit is the HERAPDF 1.0 set.

5.4. Treatment of Uncertainties

For a full understanding of the proton structure, an unbiased estimation of the uncertainties is needed. There are several different sources of uncertainty which influence the precision of the PDFs. These have to be propagated properly from their respective sources to the PDFs to finally estimate their influence on the calculated cross section.

5.4.1. Experimental Uncertainties

All measurements are affected by statistical and systematic uncertainties. The initially uncorrelated statistical uncertainty often gets correlated by propagation through an unfolding procedure, in which detector effects from finite detector resolution and acceptance are corrected. The measurement of the jet energy for example involves many corrections, which are each affected by an uncertainty. In addition, luminosity has an uncertainty. While some of the uncertainties, e.g. the luminosity uncertainty is fully correlated between all bins, other sources of uncertainty are only correlated between certain bin regions. These correlations have to be considered properly in the χ^2 minimisation.

The uncertainty on the fit parameters can be estimated by exploring the best-fit minimum through profiling the different parameters. A change in χ^2_{min} of $\Delta\chi^2 = 1$ results in a 1σ confidence level on the PDF. Since the inclusion of many datasets

with not fully understood uncertainties and correlations can lead to incompatibilities resulting in very small uncertainties on the PDFs, some PDF fitting groups, CTEQ for example, introduce a tolerance which allows the $\Delta\chi^2$ to be higher.

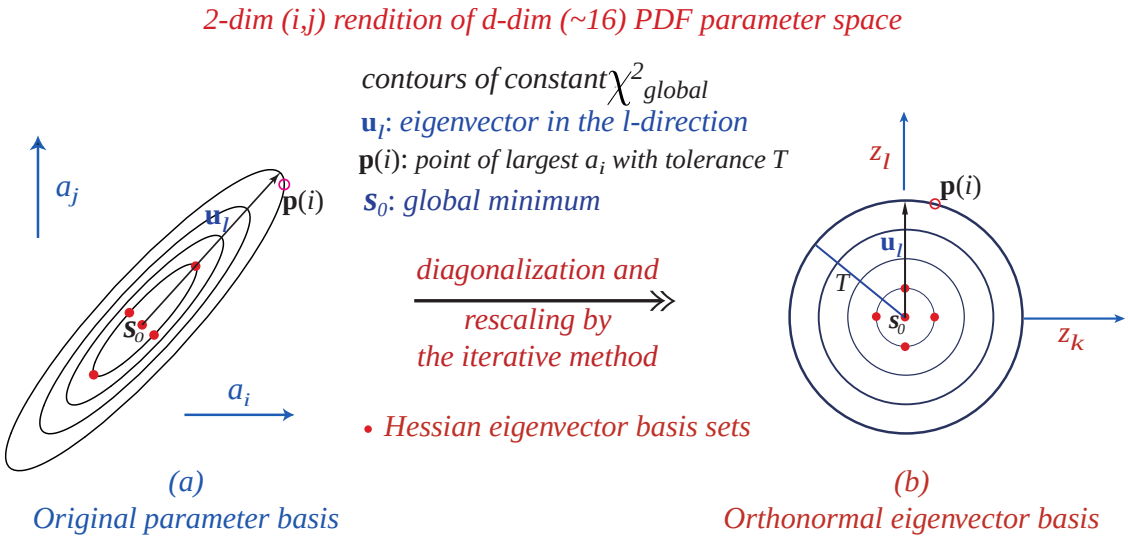


Figure 5.1.: Transformation from the original parameter basis to the orthonormal eigenvector basis. The uncertainty on the PDF can be propagated to a physical quantity X using these eigenvector PDF sets [49].

The uncertainty on the fit parameters is not very useful for propagating the PDF uncertainty, since the fit parameters are correlated leading to a non-diagonal error matrix. A better method is to do an orthogonal transformation from the parameter basis to the eigenvector basis in the parameter space. The first step is the calculation of the Hessian matrix, consisting of the second derivatives evaluated at the global minimum. The eigenvectors are derived by diagonalisation of the Hessian matrix. The eigenvectors provide a natural basis to express variations around the minimum. The procedure is shown exemplary in Figure 5.1. The impact on the best-fit χ^2 is explored by iteratively profiling the lengths of the Hessian eigenvectors. The result are eigenvector PDF sets, which cover positive and negative changes to each eigenvector. Since the eigenvectors are de-correlated, the asymmetric uncertainty on a physical quantity X can be calculated with the Equations 5.15.

$$\begin{aligned} \Delta X_+ &= \sqrt{\sum_i^{N_{\mathrm{EV}}} \left[\max(X_i^+ - X_0, X_i^- - X_0, 0)\right]^2} \\ \Delta X_- &= \sqrt{\sum_i^{N_{\mathrm{EV}}} \left[\min(X_i^+ - X_0, X_i^- - X_0, 0)\right]^2} \end{aligned} \tag{5.15}$$

parameter	nom. value	lower limit	upper limit
f_s	0.31	0.23	0.38
m_c [GeV]	1.4	1.35	1.65
m_b [GeV]	4.75	4.3	5.0
Q^2_{min} [Gev^2]	3.5	2.5	5.0

Table 5.1.: Input parameters in the fit and the variations studied following the prescription of the HERAPDF analysis [16].

5.4.2. Model Uncertainties

The determination of the PDF relies on many initial parameters. These have been varied one by one in the fit to estimate their impact on the PDFs. These parameters includes the heavy quark masses m_{c} and m_{b}, the strangeness fraction f_{s} and the minimum imposed Q^2_{min} on the data. To estimate the uncertainties, the upper and lower limits as shown in Table 5.1 have been considered in the estimation of the model uncertainties.

Each parameter is varied and the difference between the central fit and each model variation fit is added in quadrature separately for positive and negative deviations as shown in Equation 5.16. This represents the model uncertainty of the HERAPDF set.

$$\Delta X_{\pm} = \pm \sqrt{\sum_{i}^{\text{EV}} \left[\max(X_i^{\pm} - X_0, 0)\right]^2} \tag{5.16}$$

5.4.3. Parametrisation Uncertainties

The parametrisation of the PDF and the starting scale are arbitrarily chosen and changes to theses choices affect the PDF. There are several possibilities to include this uncertainty in the fit. The NNPDF collaboration uses a data-driven regularisation ansatz to estimate the influence of different parametrisations [14]. HERAPDF uses a reduced parametrisation for the fit and observes the influence on the PDF by adding more parameters to the fit. As mentioned before, the central fit includes the minimisation of 10 free parameters. For each PDF, the additional parameters D and E are added one by one and the difference to the central fit is studied. Additionally, the starting scale is treated as a parametrisation uncertainty, as the shape of the PDF strongly depends on the chosen starting scale. The starting scale Q^2_0 is varied between the lower limit $Q^2_{0,\text{low}} = 1.5\,\text{GeV}^2$ and the upper limit $Q^2_{0,\text{high}} = 2.5\,\text{GeV}^2$. The differences between all variation fits and the central fits are combined to form an envelope around the central fit and describe the parametrisation uncertainty and can be calculated using the Equation 5.17.

$$\Delta X_+ = \max_i^{\mathrm{EV}}(X_i - X_0, 0)$$
$$\Delta X_- = \max_i^{\mathrm{EV}}(X_0 - X_i, 0) \tag{5.17}$$

5.5. Sensitivity of the PDFs to Inclusive Jet Data from CMS

In this thesis, the PDF sensitivity to inclusive jet data from the CMS experiment has been studied. First tests with the HERAFitter framework have been performed using data from the 2010 LHC run. However, the statistical precision and the lack of correlated systematic uncertainties are not suited for a concluding statement on the PDF sensitivity. Hence, the study has been improved and repeated with data from the 2011 LHC run with a higher statistical precision and a carefully performed study of correlations between the systematic uncertainties.

5.5.1. CMS Inclusive Jet Cross Section 2010

The inclusive jet cross section was measured with data from the 2010 LHC run [50] at the CMS detector. The data sample corresponds to an integrated luminosity of $34\,\mathrm{pb}^{-1}$. The cross section has been measured differentially in the jet transverse momentum and the absolute values of the rapidity. To cluster the particles into jets, the Anti-k_T jet algorithm with a jet size of $R = 0.5$ has been employed. The spectrum is shown in Figure 5.2. The overall uncertainty has been estimated by quadratically adding the statistical uncertainty and uncertainties resulting from the unfolding procedure and jet energy corrections.

Unfortunately, the information about the correlation between the different sources has not been provided. Therefore, the jet energy correction uncertainties were treated as uncorrelated between the different bins in p_T and $|y|$ while the luminosity uncertainty has been treated as correlated between all bins. The treatment of the uncertainties in this way is more conservative and probably overestimates the uncertainty.

This measurement has been compared to next-to-leading-order perturbative calculations derived using the NLOJet++ cross section integrator within the framework of fastNLO. Figure 5.3 shows the fully corrected measured cross section as a ratio to the theoretical predictions. These are generally in agreement within the experimental and theoretical uncertainties. This analysis was used for an initial test with the HERAFitter framework to probe the impact of LHC jet measurements on PDF.

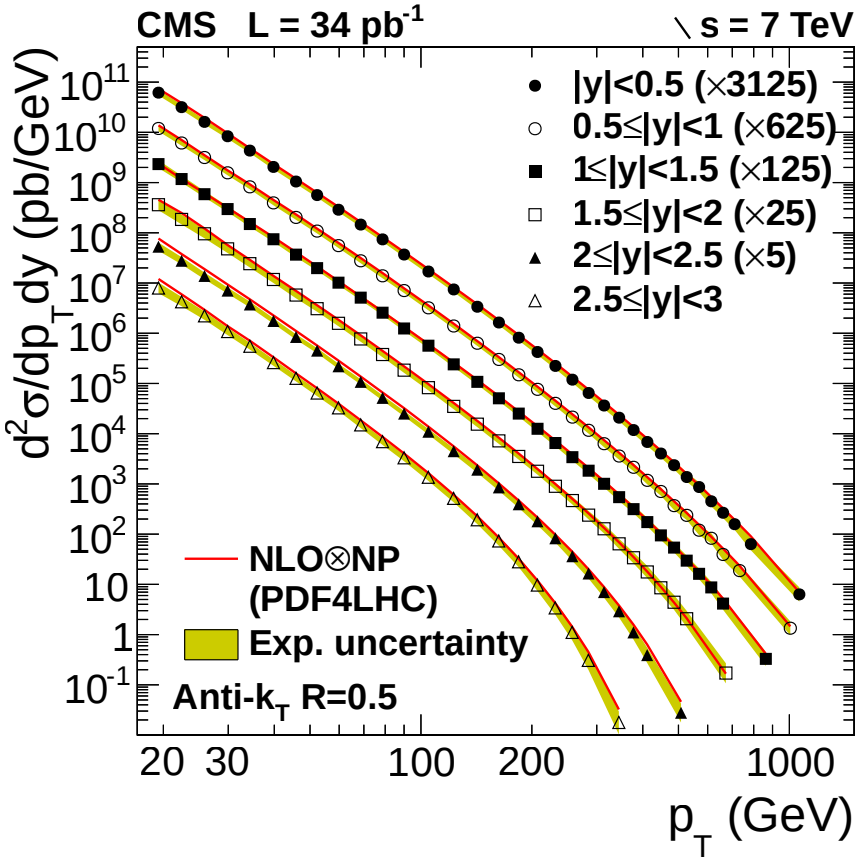


Figure 5.2.: The inclusive jet cross section measured with the CMS detector at the 2010 run of the LHC. The measured and calculated cross sections generally agree within uncertainty [50].

PDF Sensitivity

The constraints by this dataset on the PDFs and the PDF uncertainty have been studied in this analysis in a combined fit of HERA DIS data and inclusive jet data from the CMS experiment.

The χ^2 without the CMS jet data after the minimisation yields $\chi^2/n_{\text{dof}} = 575.08/582 = 0.98$. This is in agreement with the result of the HERAPDF collaboration in their publication [16]. The inclusion of the jet data leads to $\chi^2/n_{\text{dof}} = 713.63/752 = 0.94$. The error estimation has been performed following the HERAPDF 1.0 prescription as described in Section 5.3.

The PDFs have been evolved to $Q^2 = 10^4\,\text{GeV}^2$ as this corresponds more to the scale of the CMS measurement. The following plots present the results of the study. The fit with the HERA DIS data alone is always placed on the left side and the fit with the additionally considered CMS jet data on the right side. At first the gluon PDF is shown in Figure 5.4. HERA data can only constrain the gluon through QCD sum rules in the high-x region whereas the inclusive jet data directly probes the gluon PDF. This is demonstrated by the reduction of

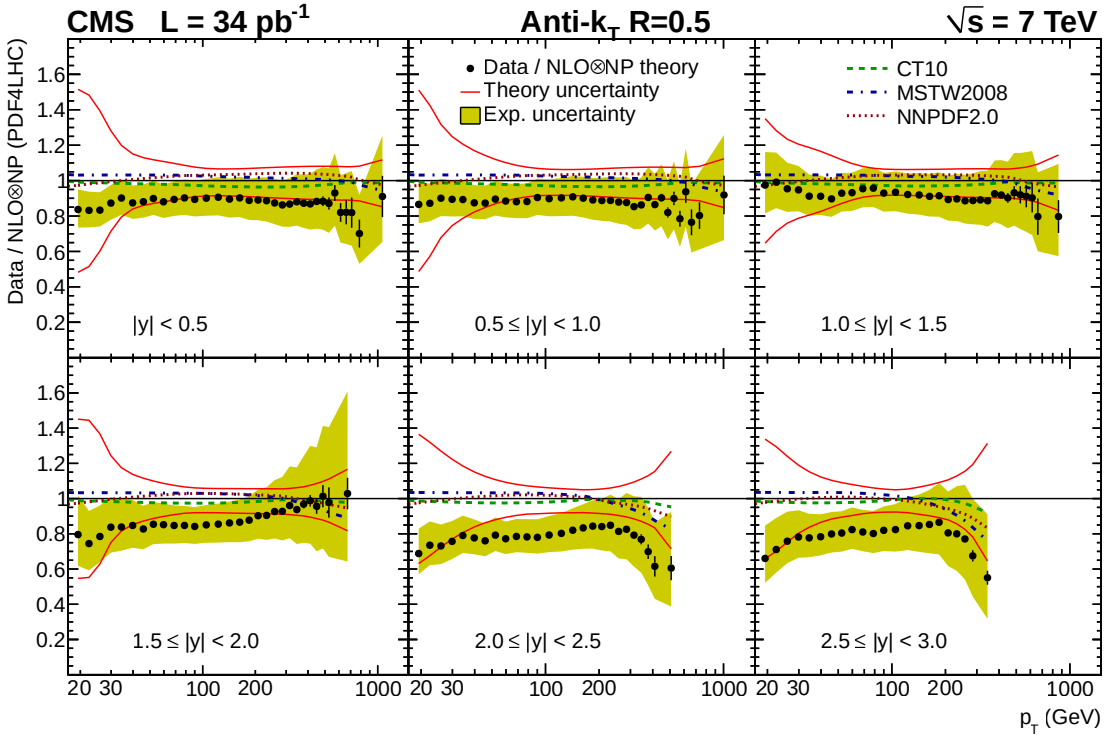


Figure 5.3.: Ratio of CMS inclusive jet 2010 measurement and predictions with three PDF sets in the different rapidity regions. The experimental and theoretical uncertainties are shown as well. These indicate a general agreement between the theory and the data [50].

Dataset	χ^2	n_{dof}
NC HERA-I H1-ZEUS combined e^-p	106.70	145
NC HERA-I H1-ZEUS combined e^+p	419.12	379
CC HERA-I H1-ZEUS combined e^-p	19.84	34
CC HERA-I H1-ZEUS combined e^+p	29.42	34

Table 5.2.: The results of a HERA-I DIS fit. The table shows the resulting χ^2 per degree of freedom for each included dataset. The total χ^2 per n_{dof} is 0.98. This PDF is very similar to the published PDF set HERAPDF 1.0.

the parametrisation uncertainties in the high-x range. The negligible effect on the experimental uncertainties can be explained with the high statistical uncertainties and the unknown systematic correlations. The effect is much better visible in fits with the more constraining 2011 data which are pres A ratio of both fits is shown in Section A.1.2.ented in Section 5.5.2.

The sub-process fraction of quark-quark processes in $2 \to 2$ scattering is dominant in the region above 800 GeV. Therefore the up- and down-quark PDFs can also be constrained by inclusive jet data. The up-quark PDF is shown in Figure 5.5,

Dataset	χ^2	n_{dof}
NC HERA-I H1-ZEUS combined e^-p	108.01	145
NC HERA-I H1-ZEUS combined e^+p	429.63	379
CC HERA-I H1-ZEUS combined e^-p	20.55	34
CC HERA-I H1-ZEUS combined e^+p	29.38	34
CMS inclusive jets 2010	119.46	170

Table 5.3.: The result of a combined fit of HERA DIS and 2010 CMS inclusive jet data. The table shows the resulting χ^2 per degree of freedom for each included dataset. The total χ^2 per n_{dof} is 0.94. Data from both experiments are compatible in a combined fit. The constraints by the jet data lead to slightly different results of the HERA DIS data compared to Table 5.2.

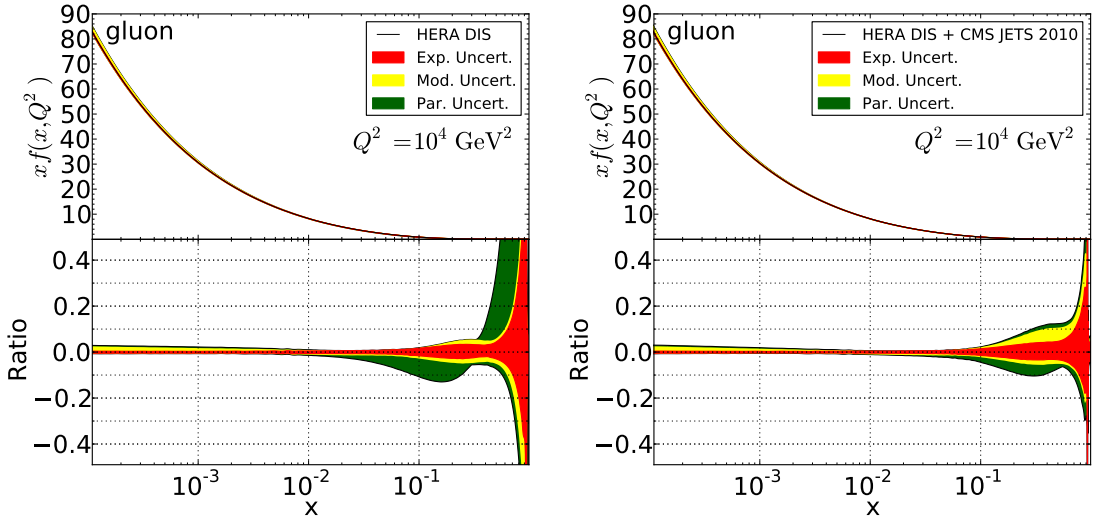


Figure 5.4.: Fit of the gluon PDF with HERA DIS data (left) and additionally inclusive jet data 2010 from CMS (right). The PDF is evolved to $Q^2 = 10^4\,\mathrm{GeV}^2$. The influence of the jet data is mainly visible by a reduction of the parametrisation uncertainties.

where the model uncertainties are reduced. The PDF of the down-quark is shown in Figure 5.6. The uncertainties of the down-quark PDF are reduced in some regions, but also slightly increase in other regions. A ratio of both fits is shown in Section A.1.2.

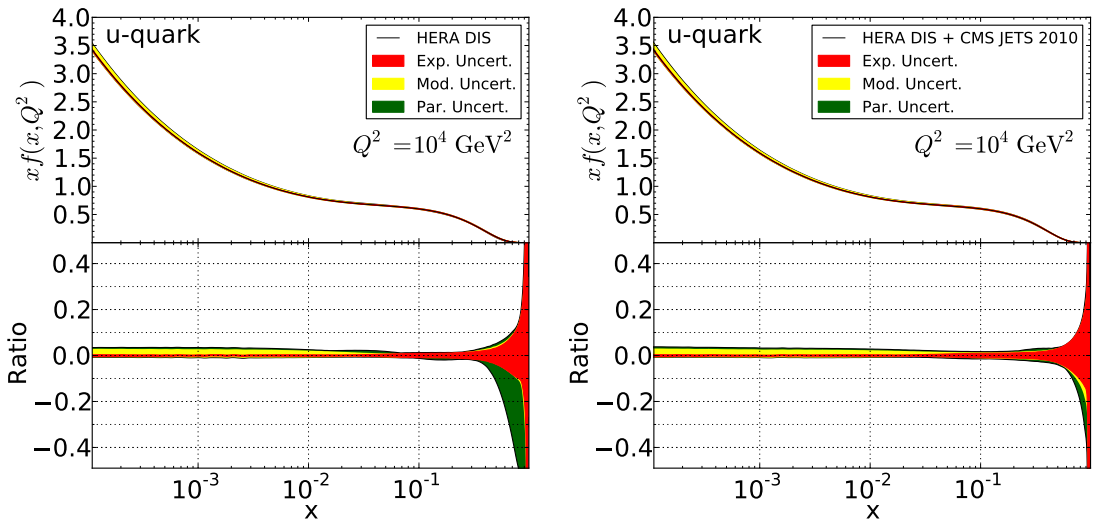


Figure 5.5.: Fit of the up-quark PDF with HERA DIS data (left) and additionally inclusive jet data 2010 from CMS (right). The PDF is evolved to $Q^2 = 10^4\,\mathrm{GeV}^2$. The influence of the jet data is mainly visible by a reduction of the parametrisation uncertainties.

The possibility of a simultaneous fit of the PDFs and the strong coupling constant has been tested. The same input assumptions and fit settings as before have been used, but the strong coupling constant has been included in the fit as an additional free parameter. Since the uncertainty on α_s has not been estimated in the fits before, the inclusion of the strong coupling allows the estimation of the α_s uncertainty. The uncertainty on α_s is evaluated together with the other parameters in the fit representing the experimental uncertainties. Therefore it is shown in the band of the experimental uncertainties. This leads to a dramatic increase of the uncertainty in a fit with only HERA DIS data as can be seen in Figure 5.7. It is shown that the inclusion of jet data significantly reduces the experimental uncertainty again, as it is able to de-correlate the gluon PDF and α_s. The result of this fit gives a rather high value of $\alpha_s(M_Z^2) = 0.1262$. The reason is not fully understood, as the theoretical prediction over-estimates the CMS data by about 10 % as can be seen in Figure 5.2. It was expected that a lower α_s value as the default $\alpha_s(M_Z^2) = 0.1176$ would be preferred, since this would result in a lower cross section prediction.

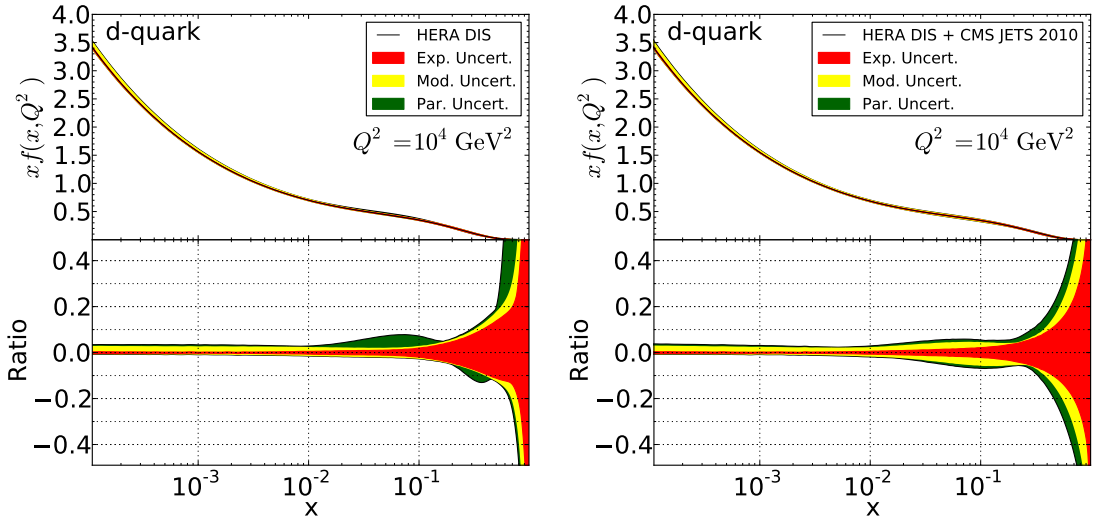


Figure 5.6.: Fit of the down-quark PDF with HERA DIS data (left) and additionally inclusive jet data 2010 from CMS (right). The PDF is evolved to $Q^2 = 10^4\,\mathrm{GeV}^2$. The influence of the jet data is mainly visible by a reduction of the parametrisation uncertainties in some regions. However there is also a slightly increase especially by the model uncertainties.

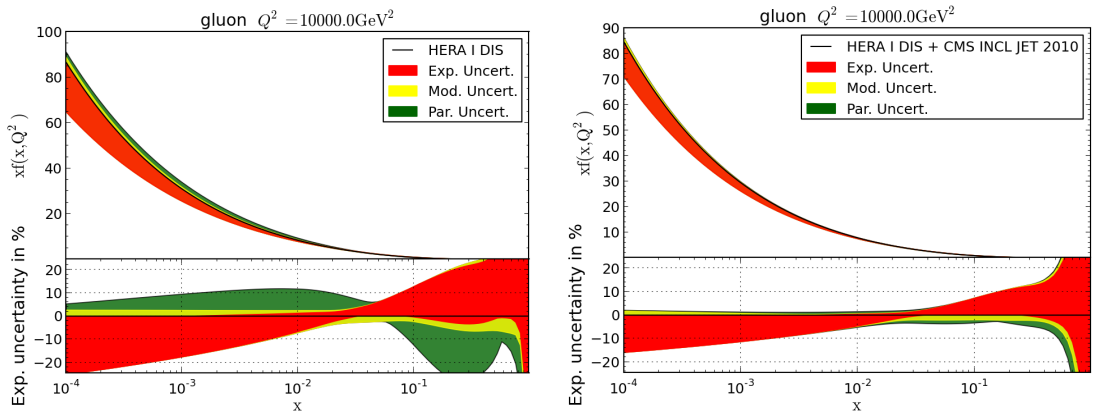


Figure 5.7.: Fit of the gluon PDF with HERA DIS data (left) and additionally inclusive jet data 2010 from CMS (right). The strong coupling α_s is included as additional free parameter in the fit. The uncertainty on α_s is embedded in the experimental band. The de-correlating effects of the inclusive jet data are visible very well. The PDF is evolved to $Q^2 = 10^4\,\mathrm{GeV}^2$.

5.5.2. CMS Inclusive Jet Cross Section 2011

Compared to the inclusive jet analysis with data from the year 2010, the new analysis includes collision data corresponding to $5.0\,\mathrm{fb}^{-1}$ of integrated luminosity. The data were collected from LHC proton-proton collisions at 7 TeV. The clustering of the jets was performed with the Anti-k_T clustering algorithm with a jet size of $R = 0.7$. The inclusive jet cross section was measured differentially in terms of the transverse momentum and the absolute rapidity. Due to the large data sample, this measurement has very small statistical uncertainties in the low-pt range and can probe QCD up to a transverse momentum of 2 TeV.

The measurement was unfolded to correct for detector effects. A detailed study of all systematic uncertainties with their correlation allows to include this measurement in PDF fits. These data are also used to extract the strong coupling in Section 5.6.

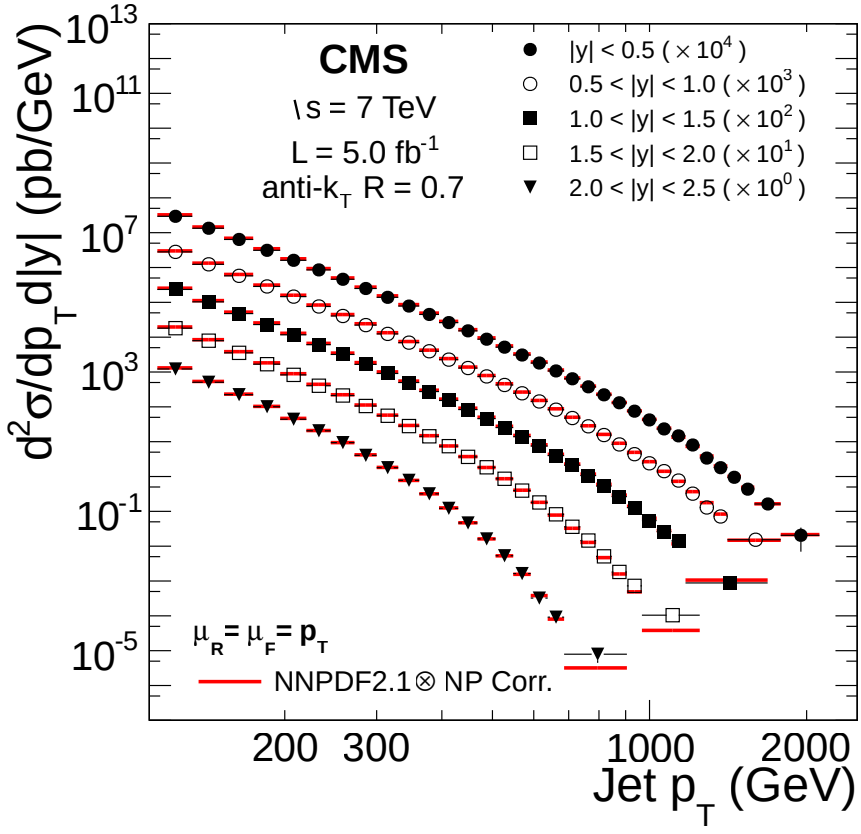


Figure 5.8.: Double differential inclusive jet cross section in p_T and $|y|$. The pQCD calculation using NLOJet++ and the NNPDF 2.1 set is shown [51].

The size of the theoretical uncertainties is of particular interest for the extraction of the strong coupling since they are the dominant source of uncertainty especially in the forward region. Figure 5.9 shows the ratio of the data to the prediction

using the NNPDF set including the overall experimental and theoretical sources of uncertainty. The PDF uncertainty was calculated according to the prediction of the NNPDF collaboration. The scale uncertainty was estimated using the common 6-point variation which will be explained in Section 5.6.1. Especially the central rapidity regions are described very well by the theoretical calculation.

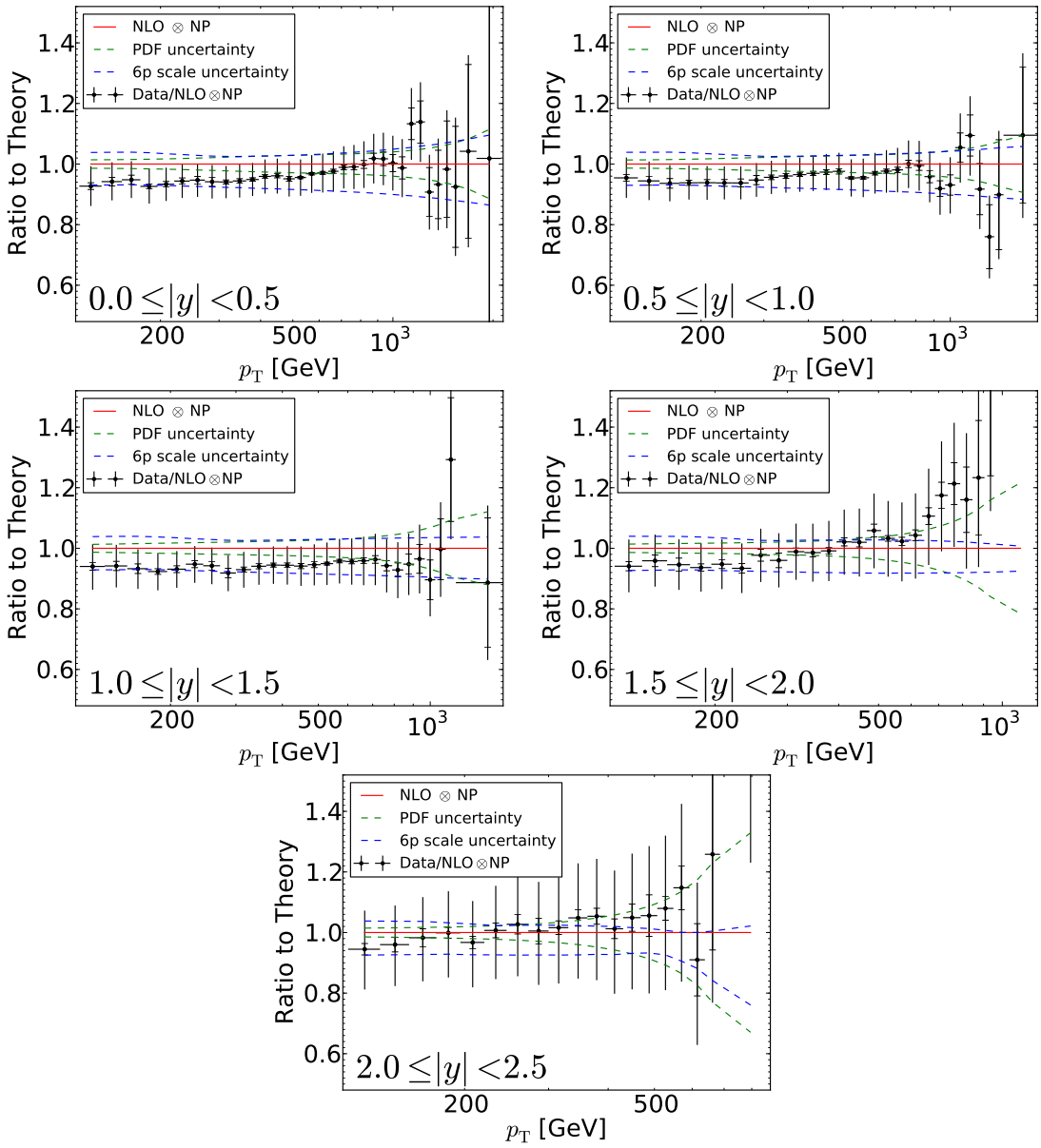


Figure 5.9.: Ratio of the measured cross section to the calculation using the NNPDF set. The error-bars show the overall experimental uncertainties. Additionally, the PDF uncertainty and the scale uncertainty are shown.

The expected impact can be shown by the correlation between the inclusive jet

cross section σ_{jet} and the PDF $i\ xf_i(x, Q^2)$. The correlation coefficient ϱ is shown in Figure 5.10 in four different phase space regions. The NNPDF collaboration provides their PDF sets as a statistical ensemble of PDF fits. These were derived by shifting the data randomly within their uncertainties and performing fits. This ensemble can be exploited statistically to gain information about the correlations. The correlation coefficient is calculated with the N_{rep} members, called replica, of the fit. The uncertainties $\Delta_{\sigma_{jet}}$ and $\Delta_{f(x,Q^2)}$ are the standard deviations of both samples. The correlation coefficient is plotted against the scaling variable x.

$$\varrho\left[\sigma_{\text{jet}}, f(x, Q^2)\right] = \frac{N_{\text{rep}}}{(N_{\text{rep}} - 1)} \frac{\langle \sigma_{\text{jet}} f(x, Q^2) \rangle - \langle \sigma_{\text{jet}} \rangle \langle x f(x, Q^2) \rangle}{\Delta_{\sigma_{\text{jet}}} \Delta_{f(x,Q^2)}} \tag{5.18}$$

In the central region the highest correlation can be observed between the gluon PDF and the inclusive jet cross section in both low-p_{T} and high-p_{T} ranges. The correlation between the quark PDFs and the cross section is quite small in this region and there should not be a huge impact visible in these PDFs. However, in the forward region, the correlation between the quark PDFs and the cross section increases in the high-p_{T} range for high values of x.

Therefore, we should see an impact on the gluon and quark PDFs in the high-x region. Unfortunately, the correlation-coefficient is not provided for the anti-quark PDFs. In any case the sub-process fraction of $q\bar{q}$ processes in $2 \to 2$ scattering is suppressed. Hence the influence on the PDFs should be small. Since the anti-quark PDFs are not well constrained by HERA DIS data, there is a non-negligible effect on the PDFs and its uncertainties still visible.

PDF sensitivity

Similar to the previous PDF study on CMS 2010 data, the same analysis strategy with some improvements has been used. The higher precision of the 2011 data should yield stronger constraints on the PDFs. The PDF analysis group CTEQ showed that considering an additional negative term in the gluon parametrisation can result in a better χ^2. Therefore, the gluon is parametrised in this study now using

$$xg(x) = A_g x^{B_g} (1 - x)^{C_g} - A_g x^{B_g} (1 - x)^{C_g}. \tag{5.19}$$

$$\tag{5.20}$$

This additional term allows the gluon to become negative in the low-x region. To ensure that only the low-x region is affected, the parameter C_g is fixed to the value 25. Nevertheless this term can lead to a negative gluon PDF in the low-x region at the starting scale. Although a negative probability distribution is of considered unphysical, this does not pose a problem here as there are no data in this region of Q^2 and x. Otherwise the constraints by the data should ensure the gluon PDF to

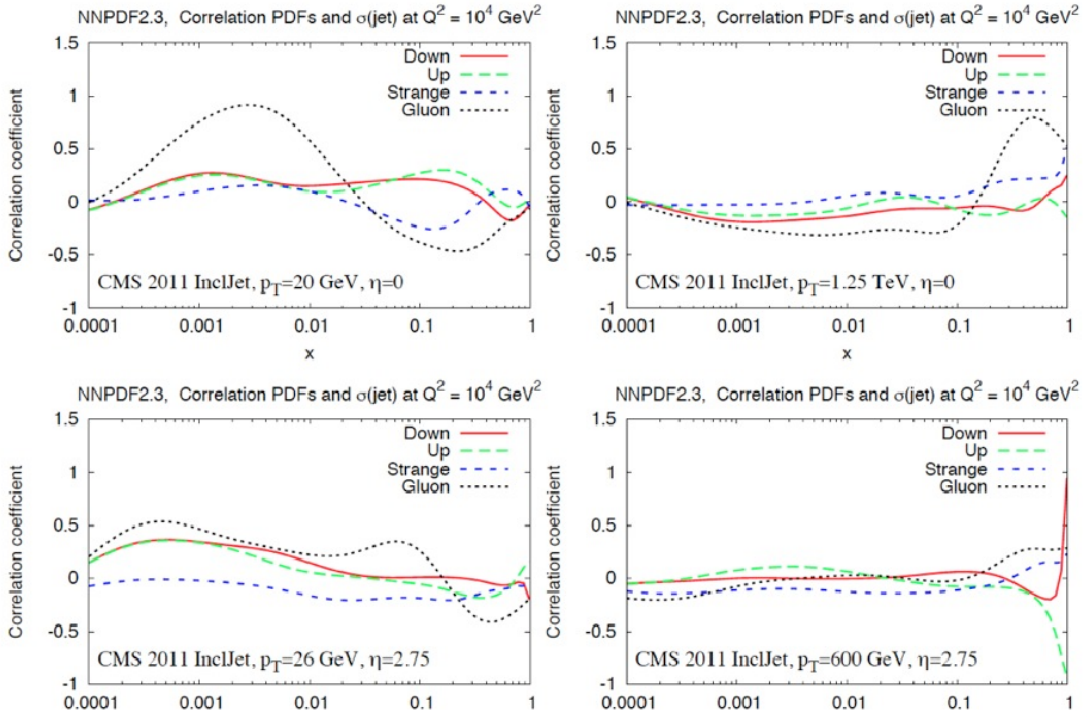


Figure 5.10.: The correlation coefficients between the cross section and the quark and gluon PDFs has been calculated in four different phase space regions. The correlation gives a hint about the expected constraints on the PDFs [52].

be positive at all times. Additionally, the constraint $B_{u_v} = B_{d_v}$ has been dropped, as it was found that this parameter is not necessary for a successful minimisation.

The results of the fit are summarised in the Table 5.4. The comparison of the results including inclusive jet data from the CMS data to the fit without show that there are no incompatibilities between these data sets. The overall fit result is χ^2/n_{dof} =756.10/711=1.06. This demonstrates that the data can be described by the fit setup. The partial χ^2/n_{dof} for the inclusive jet dataset is 162.36/133=1.22, which shows that this dataset is also well described within its uncertainties.

The uncertainties on the PDFs have been determined in the same way as in the HERAPDF 1.0 prescription. This results again in the three sources of experimental, model and parametrisation uncertainty. Similar to the last section, the PDF fits with and without jet data have been arranged next to each other to allow a direct comparison.

At first the gluon PDF is presented. Since the gluon can be probed directly by inclusive jet data, the constraints on the PDF and its uncertainty are clearly visible in Figure 5.11. The model and parametrisation uncertainties are reduced in the high-x region as expected. The uncertainty of the gluon PDF decreases in the x-region above 0.5, while there is slight increase in the x-region of 0.1 to 0.5.

Dataset	χ^2	n_{dof}
NC HERA-I H1-ZEUS combined e-p.	108.85	145
NC HERA-I H1-ZEUS combined e+p.	411.87	379
CC HERA-I H1-ZEUS combined e-p.	21.79	34
CC HERA-I H1-ZEUS combined e+p.	34.80	34
CMS inclusive jets 2011	162.36	133

Table 5.4.: The result of a combined fit of HERA DIS and CMS inclusive jet data. The table shows the resulting χ^2 per degree of freedom for each included dataset. The total χ^2 per n_{dof} is 162.36/133=1.22. Data from both experiments are compatible in a combined fit.

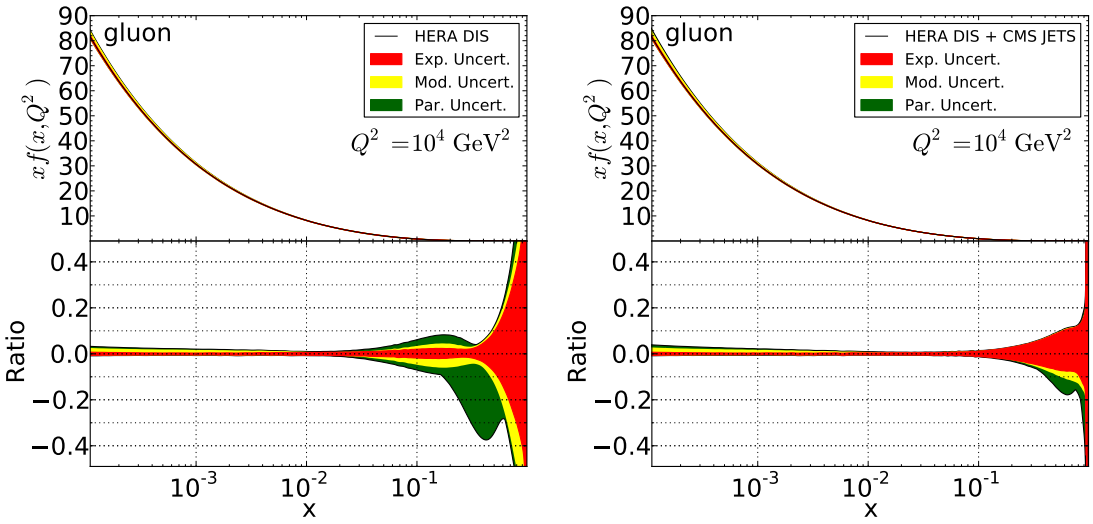


Figure 5.11.: Fit of the gluon PDF with HERA DIS data (left) and additionally inclusive jet data 2011 from CMS (right).The gluon PDF is probed directly by inclusive jet data. Due to constraints of the inclusive jet cross sections 2011, this causes a strong reduction of the model and parametrisation uncertainties as well as of the experimental uncertainty in the x-region above 0.5. The PDF is evolved to $Q^2 = 10^4\,\mathrm{GeV}^2$.

The large model and parametrisation uncertainties indicate, that this region is not constrained by the input data, but only by the chosen fit setup. The experimental uncertainties do not represent a believable estimation of the uncertainty in this region. The increase caused by the inclusion of the jet data which constrain this region, is now driven by data and the uncertainty is more believable than before. Of course, no impact is visible in the low-x region, as it is already constrained by the DGLAP evolution.

The PDFs of the up- and down-quark distribution are shown in Figure 5.12 and Figure 5.13. As indicated by the correlation plot, these can be probed by the qq sub-process in the jet production. A reduction of model and experimental uncertainties can be observed for the up-quark in the x-range close to 1. This confirms the expectations.

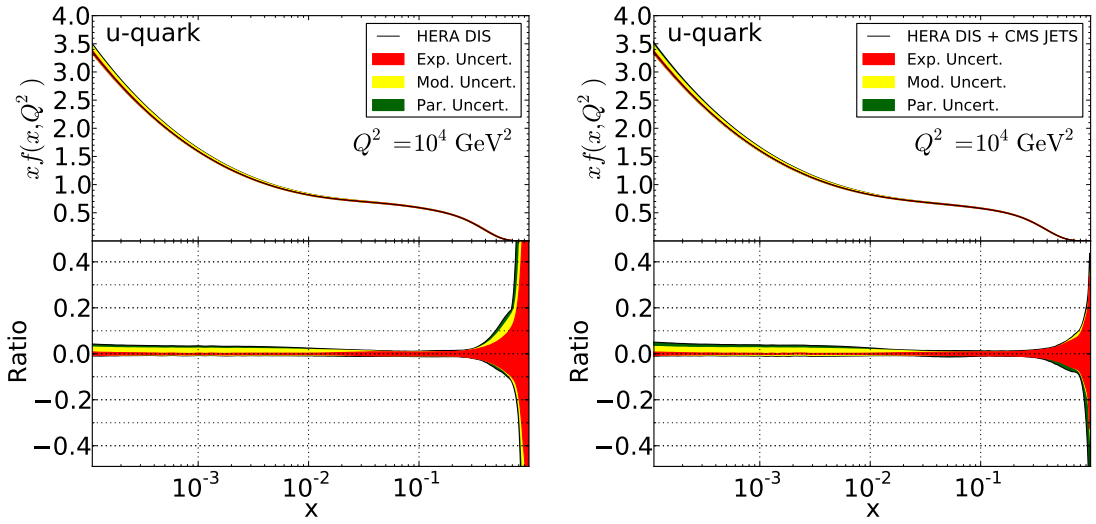


Figure 5.12.: Fit of the up-quark PDF with HERA DIS data (left) and additionally inclusive jet data 2010 from CMS (right). The qq processes in the inclusive jet production constrain the PDF in the high-x region. A reduction of the model and experimental uncertainties is visible. The PDF is evolved to $Q^2 = 10^4\,\mathrm{GeV}^2$.

The result is more ambiguous for the down-quark PDF. While the overall uncertainty is reduced as well, there is a small increase of the experimental uncertainty. This could be caused by incompatible predictions by DIS and inclusive jet data. The model uncertainties are strongly reduced due to the constraints by the jet data. This leads to an overall reduction of the uncertainty. A direct comparison of both fits can be found as ratio in the Section A.1.2.

The jet energy corrections (JEC) are the dominant source of the experimental systematic uncertainty. The steep slope of the spectrum leads to a large uncertainty in the cross section even if the p_T-scale is affected by only a small uncertainty. The sources of uncertainty taken into account for the p_T and η dependence are split into 16 uncorrelated sources which are fully correlated between themselves.

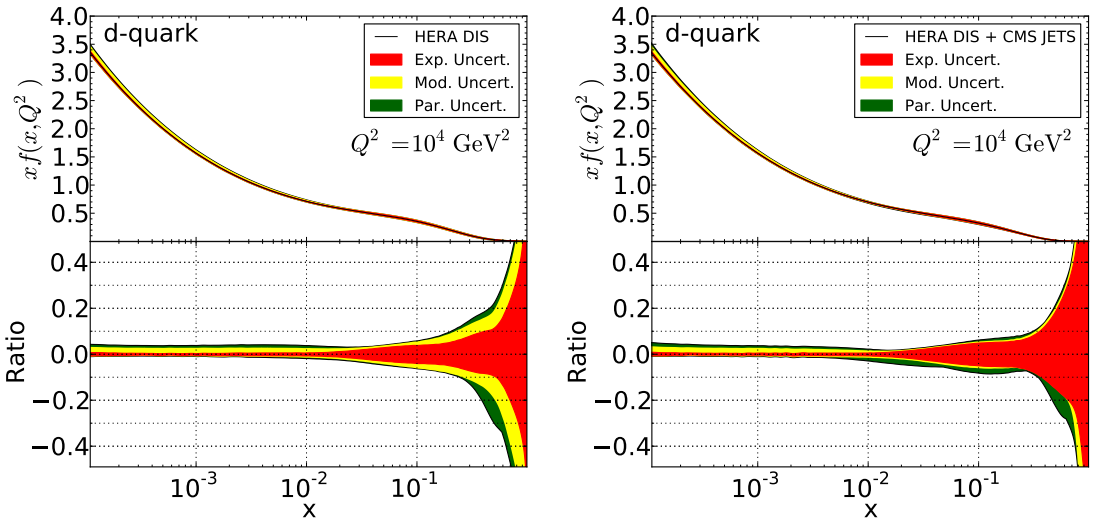


Figure 5.13.: Fit of the down-quark PDF with HERA DIS data (left) and additionally inclusive jet data 2010 from CMS (right). The qq processes in the inclusive jet production allow to constrain the PDF in the high-x region. An overall reduction of the uncertainty mainly in the model uncertainty can be observed. The PDF is evolved to $Q^2 = 10^4\,\mathrm{GeV}^2$.

These 16 sources can be split into the four categories of pileup effects, the relative calibration of the jet energy scale versus η, the absolute energy scale and differences in quark- and gluon-initiated jets. The JEC corrections are discussed in more detail in the inclusive jet analysis paper [51]. Furthermore the derivation of the specific corrections with their uncertainties is documented in [53, 54].

Because of the definition chosen for the χ^2 it is possible to observe the shift on the cross section prediction caused by each source. This is done using nuisance parameters. These give the shift of each source in units of standard deviations. The nuisance parameters are given in Table 5.5. Few of these sources should cause shifts of more than one sigma. Most of the sources do not result in a substantial shift while three of the sources cause shifts by more than one sigma.

While most of the shifts are very reasonable, the nuisance parameter of the source RelativeJERC1 causes a shift of about 2.5 sigma. This source of uncertainty is one of the eta-dependent sources derived from the jet p_T resolution. These are assumed to be fully correlated for the endcaps region within the tracking system (RelativeJERC1), the endcaps outside the tracking region (RelativeJERC2), and the hadronic forward calorimeter region (RelativeJERCHF). It is possible that the assumed correlation of the endcap region within the tracking system is overestimated resulting in a large shift of the nuisance parameter.

Systematic source	Shift in σ	Systematic source	Shift in σ
NPCORR	-0.4176	RelativeJERC2	0.0659
LUMI	-0.5254	RelativeJERCHF	0.0000
SYST UNF	1.2614	RelativeStatEC2	0.0659
ABSOLUTE	0.1934	RelativeStatHF	0.0000
HIGH PT EXTRAPOL.	0.1394	RelativeFSR	1.3532
SINGLE PION	-0.6945	PileUpDataMC	0.3463
FLAVOUR	0.1537	PileUpOOT	0.3286
TIME	-0.0935	PileUpPt	0.0417
RelativeJERC1	-2.5769	PileUpBias	-0.2704
PileUpJetRate	0.4773		

Table 5.5.: The inclusive jet analysers determined 19 independent systematic uncertainties. The shift of each systematic uncertainty is determined by nuisance parameters in the fit. The shift of each source of uncertainty on the observable in units of the uncertainty can be observed directly.

5.6. Determination of the Strong Coupling from Inclusive Jet Cross Sections

Jet production at the LHC is the dominant process and jet measurements probe QCD at high scales. The inclusive jet production depends on the value of the strong coupling and allows to extract α_s. Due to reduced statistical and precisely studied systematic sources of uncertainty, the strong coupling could be extracted from the 2011 data set.

There are two possible strategies for an extraction of α_s from cross section measurements. The approach based on the less input assumptions while treating all correlation correctly is a simultaneous fit of the PDFs and the strong coupling. This allows to decouple the strongly correlated gluon parameters from the strong coupling constant. However, this method showed systematic biases due to the impact of the relatively large normalisation uncertainties, which allow α_s to shift freely in a wide range while the changes on the cross section were compensated by the gluon parametrisation and the normalisation uncertainty. The complexity of this approach did not permit to include this into the presented study. This will be investigated in the near future.

The other possibility is to use PDFs provided by the global PDF fitting groups. These fitting groups perform PDF fits for a series of fixed values of α_s. These series of PDFs is employed to calculate cross sections and compare them to the measurement. This has the advantage to benefit from the carefully determined global PDF fits, which include data from many experiments to determine the PDFs

with high precision. Out of these global PDF fits, NNPDF has been chosen for this study for two reasons. It provides a statistically very stable and unbiased determination of the uncertainties and it provides this uncertainty for each α_s-variation. The other PDF fitting groups only provide PDF uncertainties for their central α_s PDF set. The prediction of the inclusive jet cross section was calculated for each α_s-variation PDF set, which are available in a range from 0.106 to 0.124 in steps of 0.001. The ratio of each prediction to the central PDF prediction (with $\alpha_s = 0.119$) is shown in Figure 5.14. Ratio plots for the PDF sets CT10, ABM11 and MSTW2008 are shown in Section A.2.1. The measured cross section is covered by these α_s-variations in a wide range of the phase space. The comparison between data and these predictions can be employed to extract the strong coupling.

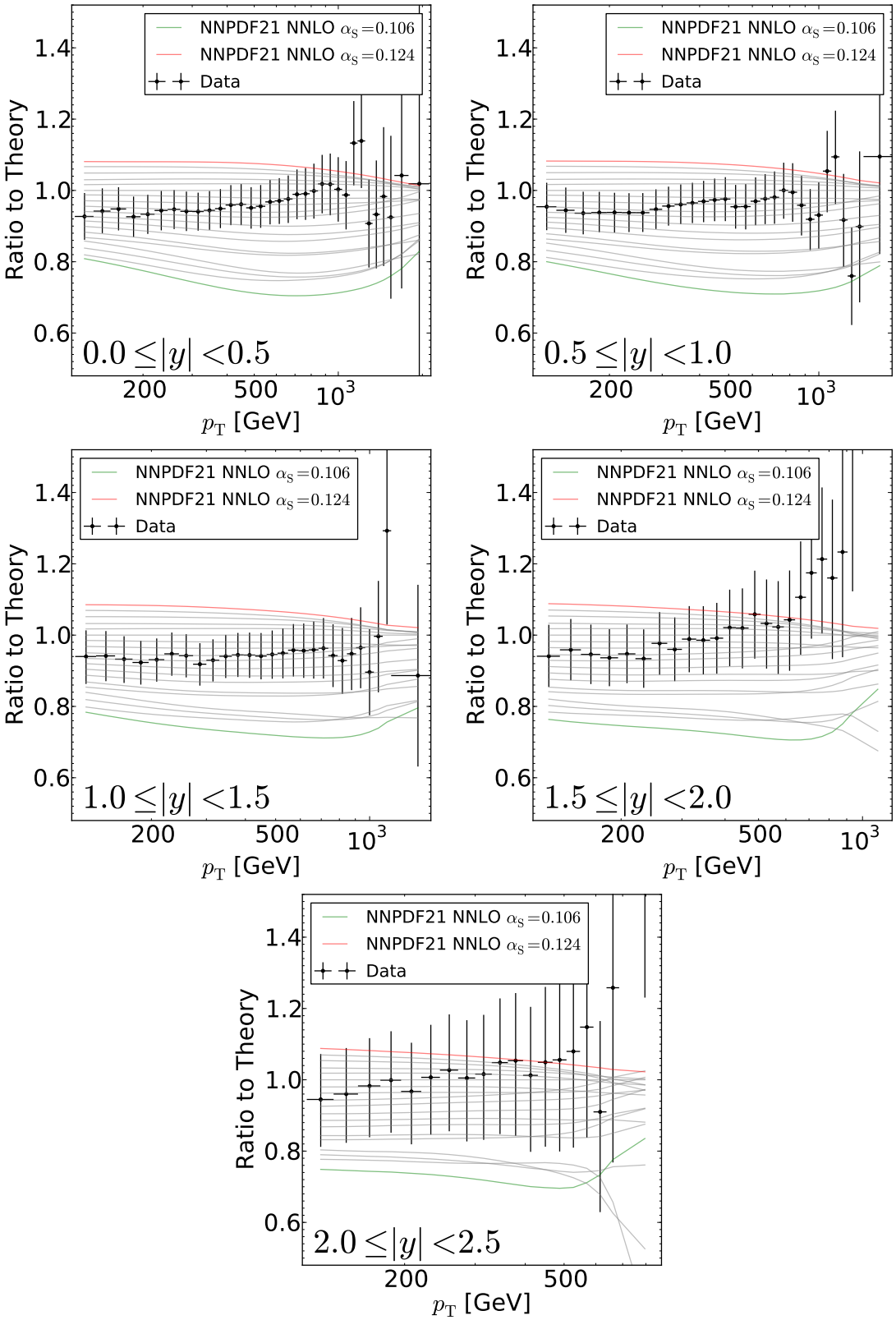


Figure 5.14.: Ratio of the cross section prediction using all α_s-variation fits to the standard fit. The fits for fixed values of α_s are available in a range of 0.106 to 0.124 in steps of 0.001. The data points show the systematic and statistical uncertainties added in quadrature and shown as error bars.

5.6.1. The Method

Similar to the method used in the PDF determination, a χ^2-estimator is used. As the nuisance parameter of the systematic uncertainties are not so important in these kind of fits and the computation time is negligible compared to full PDF fits, a standard χ^2-method using covariance matrices has been used.

The covariance matrix of each source of uncertainty has been calculated while considering the known correlations between the bins. All covariance matrices have been added to get the resulting covariance matrix for the fit.

$$\text{cov} = \text{cov}^{\text{stat}} + \sum \text{cov}^{\text{JES}} + \text{cov}^{\text{unf}} + \text{cov}^{\text{uncorr}} + \text{cov}^{\text{NP}} + \text{cov}^{PDF} \tag{5.21}$$

Only the scale uncertainty has not been included in the fit, as it does not represent an uncertainty in terms of a standard deviation. It is a convention of how to consider influences by higher orders in pQCD. It is common use to estimate the scale uncertainty by separately changing the renormalisation and the factorisation scale independently by factors of 1/2 and 2. The two variations $\mu_r/\mu_f = 0.5/2.0$ and $\mu_r/\mu_f = 2.0/0.5$ are not taken into account [55]. The fits of the strong coupling have been done for each of these six scale variations. The maximum deviation of α_s to the central value of α_s describes the scale uncertainty, which is then reported separately.

The χ^2 function as used in the fits is

$$\chi^2 = \sum_{ij}^{N} (D_i - T_i) V_{ij}^{-1} (D_j - T_j) \tag{5.22}$$

with the residual vector $(D_i - T_i)$.

All sources of systematic uncertainty have been treated as multiplicative by re-scaling them by the ratio of T/D. This ensures that the result is not affected by the d'Agostini bias, which would lead to lower values of α_s.

Since there are only PDF fits with discrete values of α_s available, the χ^2-test is done for each of the α_s-variation fits and the resulting χ^2 values are parametrised using a polynomial function. The central α_s value is obtained by the minimum of the χ^2-parametrisation. The uncertainty is derived by a profile likelihood in α_s, where the 68 % CL defines the value for which the χ^2 is increased by 1 compared to the minimum value. The separate uncertainty contributions can be estimated by removing them one by one and redoing the fit study.

5.6.2. Evaluation of $\alpha_s(M_Z^2)$

The strong coupling has been extracted separately for each rapidity bin resulting in five α_s values for the different rapidity regions. A combined fit of all rapidity regions is possible but it is non-trivial to consider all correlations between the

different regions correctly. If the correlation between the rapidity bins is overestimated, a bias to the result is introduced. This combined fit is an ongoing study. The five different results are, of course, correlated and do not represent completely independent determinations.

Table 5.6 shows the accessible p_{T} region of each rapidity bin together with the extracted value of $\alpha_{\mathrm{s}}(M_Z^2)$. All α_{s} values and their uncertainties are evolved to the scale of the Z-boson mass, as it is common usage to present $\alpha_{\mathrm{s}}(M_Z^2)$. The total uncertainty has been split into three sources. The first one shows the evaluated experimental uncertainty from the fit, the second the uncertainty due to the PDFs and the third one shows the scale uncertainty. The goodness of the fit is estimated by χ^2/n_{dof}. as expected, the uncertainty on α_{s} is smaller in the central region and increases in the forward region. The reduced χ^2 is often somewhat smaller than 1. This indicates an overestimation of the uncertainties. The probably not completely correct estimated correlations between the $|y|$ bins lead to an increased reduced χ^2 slightly larger than 1 in fits combining all rapidity bins. The plots for all six scale variations are shown in Section A.2.2.

p_{T} range	$\alpha_{\mathrm{s}}(M_Z^2)$	Exp	PDF	Scale	χ^2/n_{dof}
114 - 2116	0.1159	$^{+0.0016}_{-0.0015}$	$^{+0.0018}_{-0.0017}$	$^{+0.0018}_{-0.0009}$	16.65/33=0.50
114 - 1784	0.1170	$^{+0.0017}_{-0.0016}$	$^{+0.0017}_{-0.0018}$	$^{+0.0027}_{-0.0008}$	23.62/30=0.79
114 - 1684	0.1158	$^{+0.0020}_{-0.0019}$	$^{+0.0015}_{-0.0016}$	$^{+0.0023}_{-0.0013}$	10.53/27=0.38
114 - 1248	0.1125	$^{+0.0025}_{-0.0025}$	$^{+0.0023}_{-0.0022}$	$^{+0.0027}_{-0.0018}$	21.21/24=0.88
114 - 905	0.1154	$^{+0.0044}_{-0.0036}$	$^{+0.0033}_{-0.0029}$	$^{+0.0028}_{-0.0027}$	9.46/19=0.50

Table 5.6.: The results of the α_{s} extraction for the five rapidity bins. The uncertainty is split into an experimental, PDF and scale uncertainty estimated by a 6p-variation. All results are compatible to each other and are also in agreement within uncertainties to the world average of α_{s}.

The world average of the strong coupling constant is $\alpha_{\mathrm{s}}(M_Z^2) = 0.1184 \pm 0.0007$. The results of this study are systematically lower than this value while still compatible within uncertainties. The systematically lower values are observed also in other jet measurements. DØ extracted α_{s} from inclusive jet cross sections and found $\alpha_{\mathrm{s}}(M_Z^2) = 0.1161^{+0.0041}_{-0.0048}$ [56]. The analysis of the 3-jet mass cross section from the CMS experiment and the extraction of α_{s} leads to $\alpha_{\mathrm{s}}(M_Z^2) = 0.1169^{+0.0050}_{-0.0032}$ [57]. The determination of the ratio of 3/2-jet rate from CMS also found the compatible results $\alpha_{\mathrm{s}}(M_Z^2) = 0.1143^{+0.0083}_{-0.0067}$ [58].

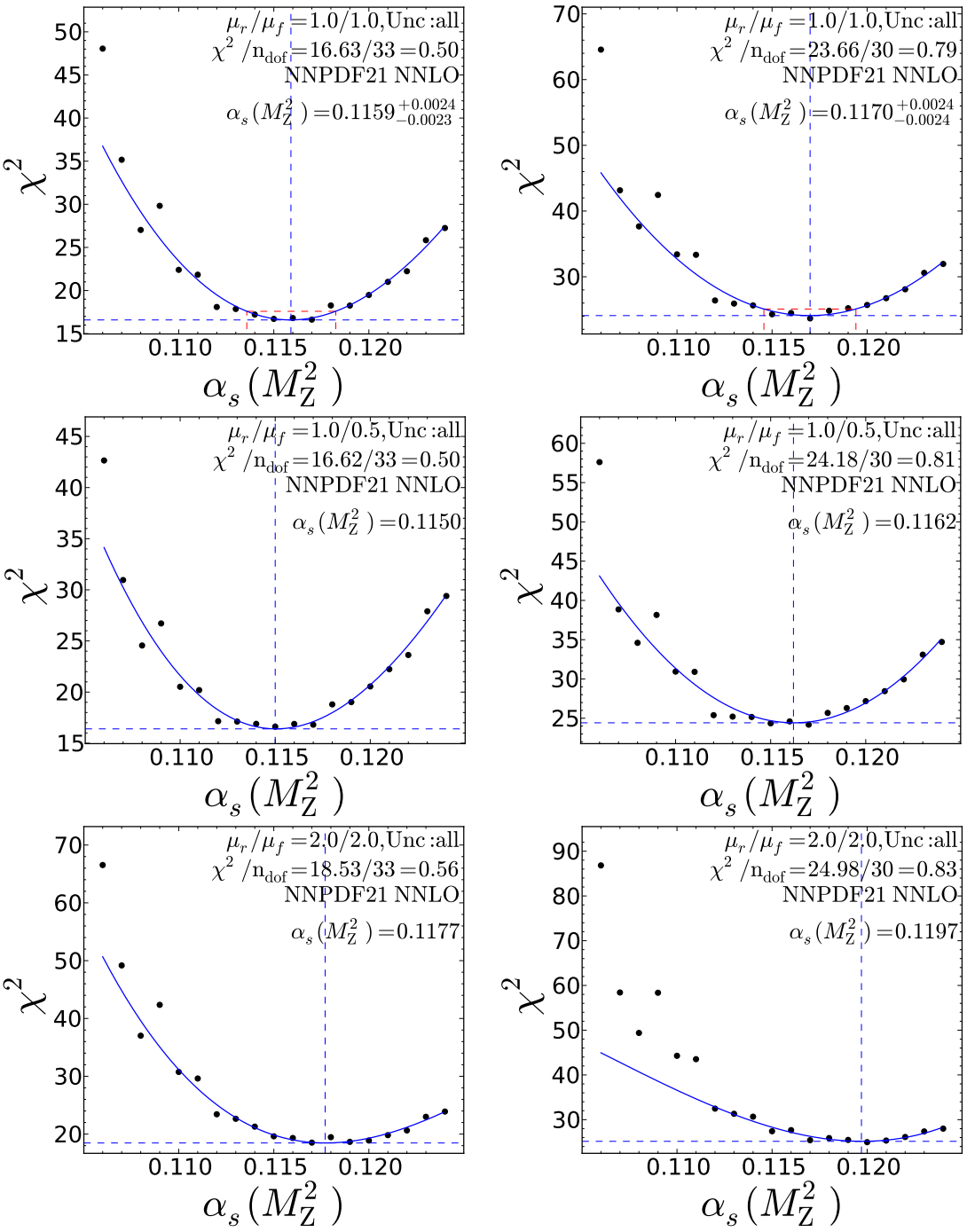


Figure 5.15.: The plots on the left side show the α_s fits of the rapidity region $\mathbf{0.0 < |y| < 0.5}$ while the plots on the right side show the fit of the region $\mathbf{0.5 < |y| < 1.0}$. The upper plots show fits with the default scale $(\mu_r/\mu_f) = (1.0/1.0)$ while the plots in the midst show the fits with the maximal downward deviations on α_s obtained by the scale variations $(\mu_r/\mu_f) = (1.0/0.5)$ and $(\mu_r/\mu_f) = (1.0/0.5)$. The deviation to the highest values of α_s shown at the bottom is always observed with the scale variation $(\mu_r/\mu_f) = (2.0/2.0)$.

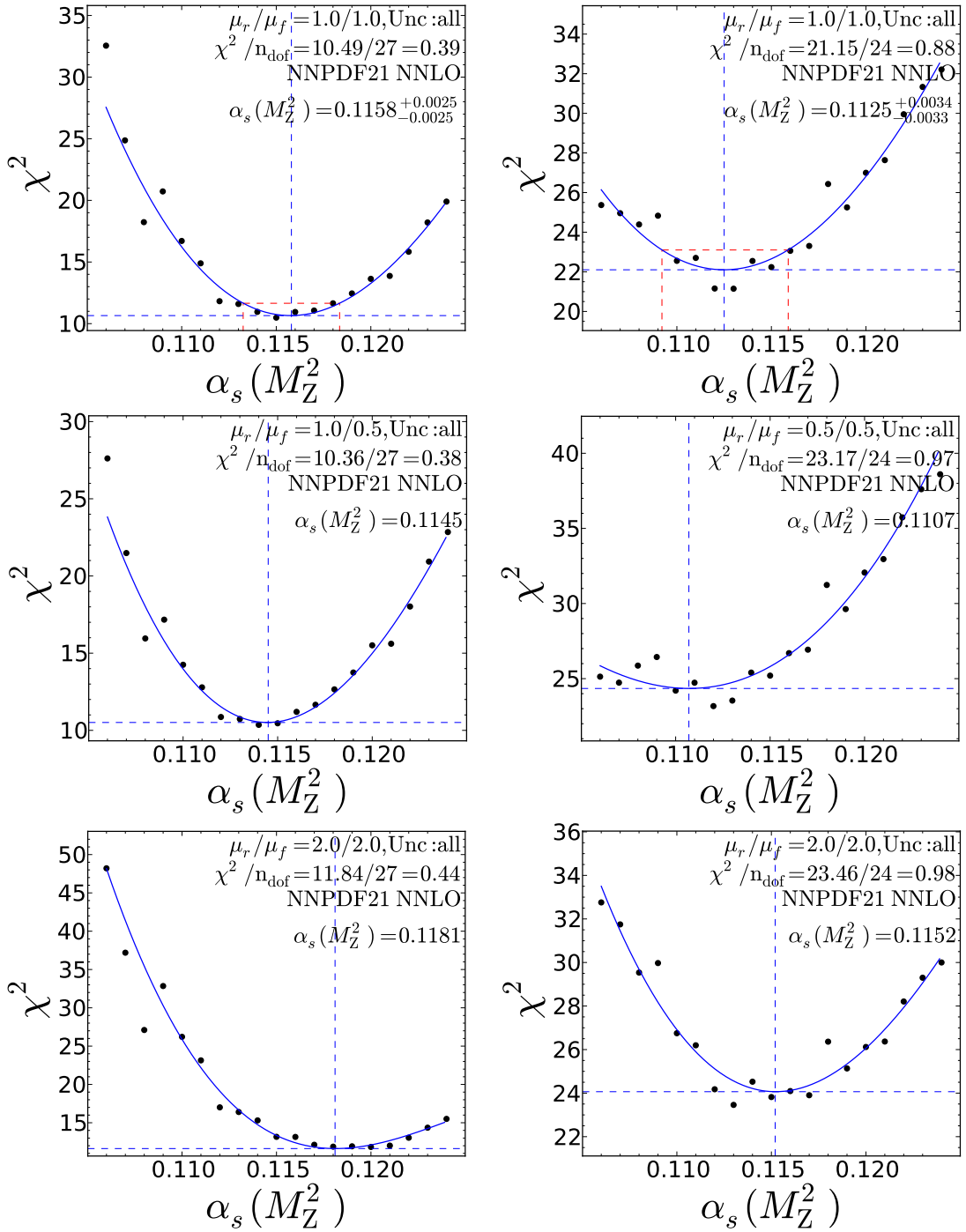


Figure 5.16.: The plots on the left side show the α_s fits of the rapidity region $\mathbf{1.0 < |y| < 1.5}$ while the plots on the right side show the fit of the region $\mathbf{1.5 < |y| < 2.0}$. The upper plots show fits with the default scale $(\mu_r/\mu_f) = (1.0/1.0)$ while the plots in the midst show the fits with the maximal downward deviations on α_s obtained by the scale variations $(\mu_r/\mu_f) = (1.0/0.5)$ and $(\mu_r/\mu_f) = (0.5/0.5)$. The deviation to the highest values of α_s shown at the bottom is always observed with the scale variation $(\mu_r/\mu_f) = (2.0/2.0)$.

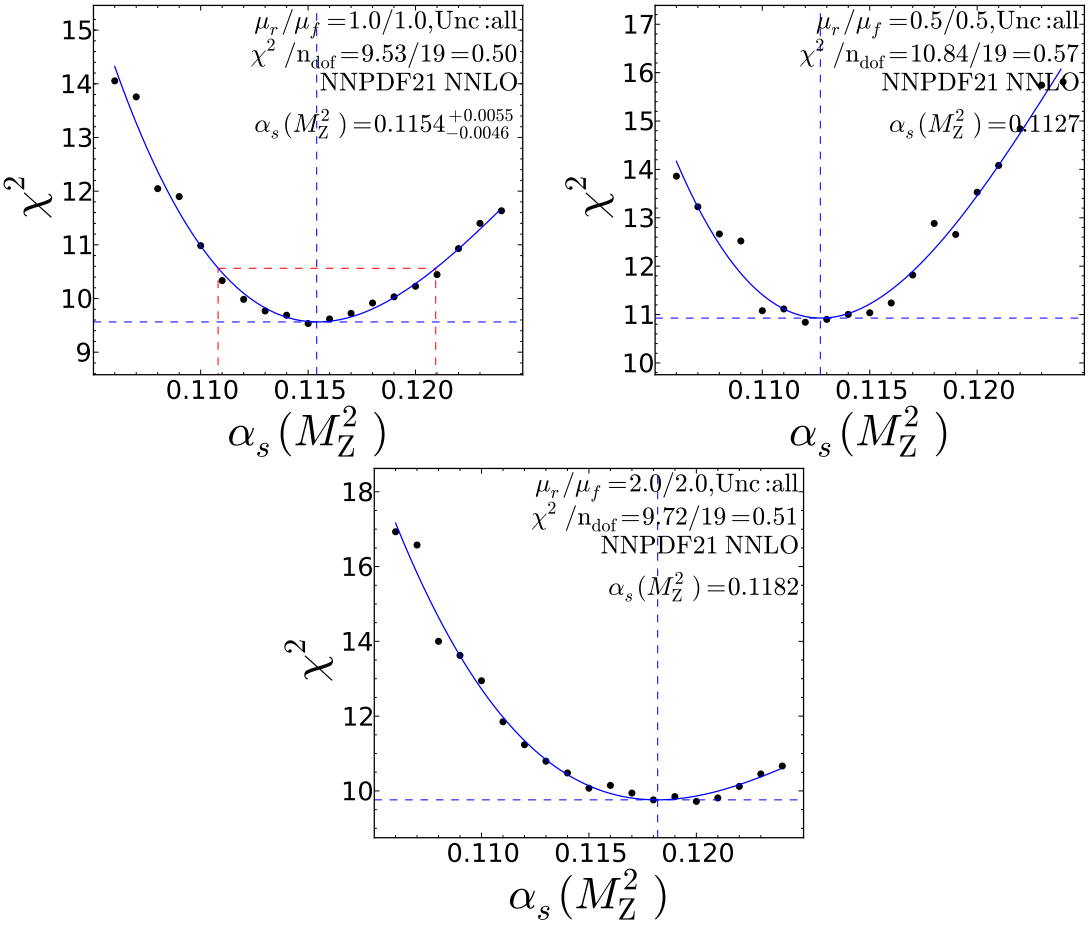


Figure 5.17.: The plots show the α_s fits for the rapidity region $\mathbf{2.0 < |y| < 2.5}$. The upper left plot shows a fit with the default scale $(\mu_r/\mu_f) = (1.0/1.0)$ while the plots on the upper right side shows the fit with the maximal downward deviation on α_s obtained by the scale variation $(\mu_r/\mu_f) = (0.5/0.5)$. The plot in the midst show the fit with the maximal upwards variation obtained by the scale variation $(\mu_r/\mu_f) = (2.0/2.0)$.

5.6.3. The Running of α_s

The running of the strong coupling predicted by the RGE evolution equation can be compared to extracted α_s values from different scales of the measurement. The central rapidity bin of the inclusive jet measurement has been studied. The p_T spectrum was divided into three regions $114\,\text{GeV} \leq p_T < 300\,\text{GeV}$, $300\,\text{GeV} \leq p_T < 737\,\text{GeV}$ and $737\,\text{GeV} \leq p_T < 2116\,\text{GeV}$. These ranges were chosen to provide approximately the same statistical precision in each region. An α_s

extraction according to the method described in the section before was performed. The $\alpha_s(M_Z^2)$ results are shown in Table 5.7. They are compatible with the world average value. The cross section weighted average of the transverse momentum in the bins of each fit range from a NLO calculation has been chosen for the determination of the scale. The evolution of α_s to the scale of the measurement has been performed using the α_s evolution code GRV [59] provided within fastNLOReader. The Table 5.8 shows the value of the strong coupling at the scale of measurement.

p_T range [GeV]	$\alpha_s(M_Z^2)$	Exp.	PDF	Scale	χ^2/n_{dof}
114 - 300	0.1163	$^{+0.0025}_{-0.0025}$	$^{+0.0011}_{-0.0010}$	$^{+0.0038}_{-0.0013}$	1.95/8=0.24
300 - 737	0.1154	$^{+0.0024}_{-0.0024}$	$^{+0.0010}_{-0.0007}$	$^{+0.0032}_{-0.0002}$	2.92/11=0.27
737 - 2116	0.1176	$^{+0.0025}_{-0.0025}$	$^{+0.0015}_{-0.0015}$	$^{+0.0029}_{-0.0008}$	11.71/14=0.84

Table 5.7.: The result of the fits evaluated at the mass of the Z-boson. Within its uncertainties the extracted α_s values are compatible with the value of the α_s world average.

It should be noted again, that these results must be treated as correlated. Figure 5.20 shows the world average value of the strong coupling α_s as a function of the momentum transfer Q compared to various measurements which were used for the determination of the strong coupling. Additionally, the results of the α_s extraction from inclusive jet cross sections are shown.

p_T range [GeV]	Q [GeV]	$\alpha_s(Q^2)$	PDF	Exp.	Scale unc.
114 - 300	142.21	0.1096	$^{+0.0025}_{-0.0026}$	$^{+0.0012}_{-0.0010}$	$^{+0.0032}_{-0.0013}$
300 - 737	392.98	0.0952	$^{+0.0016}_{-0.0017}$	$^{+0.0012}_{-0.0010}$	$^{+0.0020}_{-0.0006}$
737 - 2116	841.17	0.0884	$^{+0.0015}_{-0.0016}$	$^{+0.0015}_{-0.0010}$	$^{+0.0017}_{-0.0006}$

Table 5.8.: The extracted α_s value evolved to the weighted-average scale of the fitted p_T-range. The running of the strong coupling can be studied by evolving the world average α_s to higher scales and compare it to the extracted α_s values.

The extracted values of the strong coupling in the different scale regions are compatible with values predicted by the RGE. This demonstrates that the assumption

of an α_s evolution following the normal RGEs even at scales beyond previous experiments is consistent with the inclusive jet cross section measured by CMS.

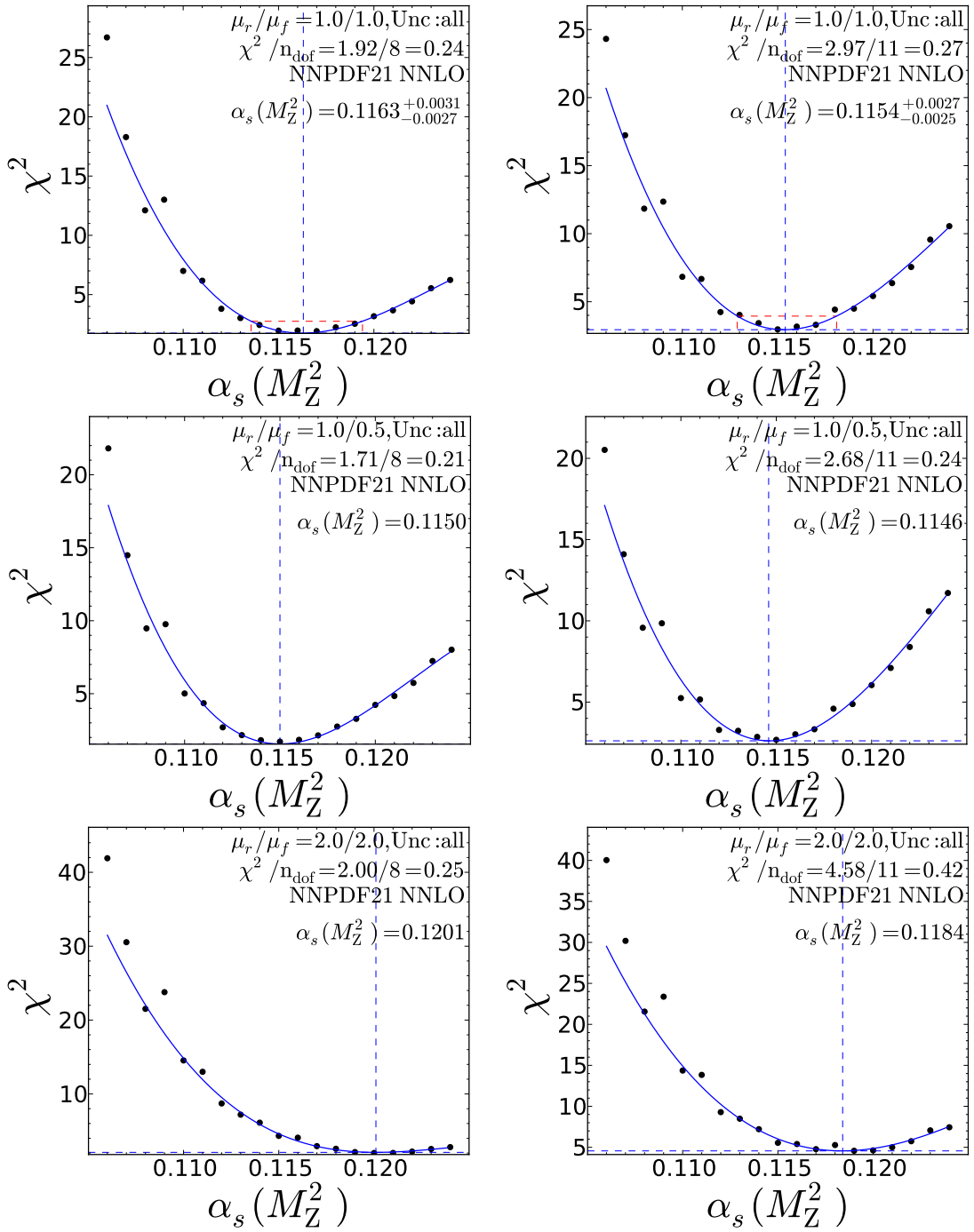


Figure 5.18.: α_s fit of the central rapidity region $0.0 < |y| < 0.5$. The p_T spectrum is split in three intervals. The left plots show the fit for the p_T region 114 GeV to 300 GeV, the right plots the p_T region 300 GeV to 737 GeV. The upper plots show fits with the default scale $(\mu_r/\mu_f) = (1.0/1.0)$ while the plots in the midst show the fits with the maximal downward deviations on α_s by the scale variations of $(\mu_r/\mu_f) = (1.0/0.5)$ and $(\mu_r/\mu_f) = (1.0/0.5)$. The deviations to the highest values of α_s shown at the bottom are always observed with the scale variation $(\mu_r/\mu_f) = (2.0/2.0)$.

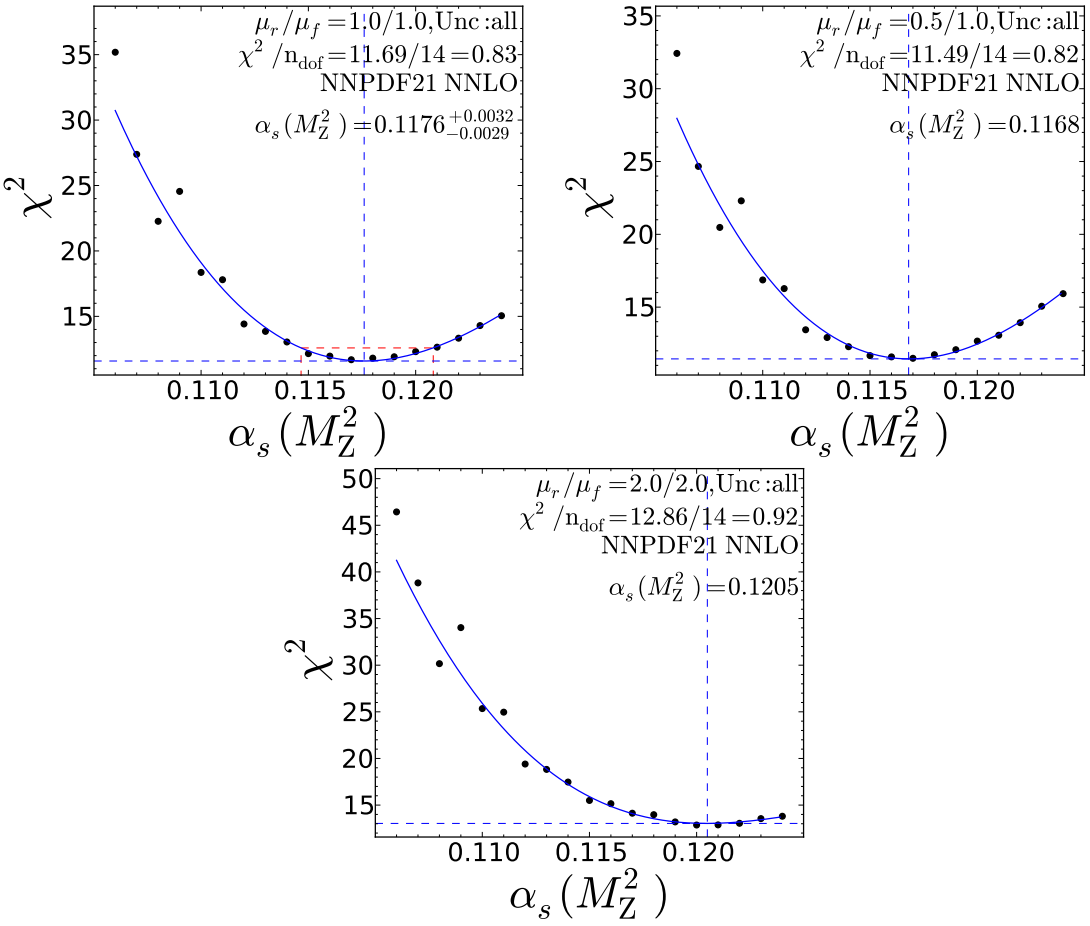


Figure 5.19.: α_s fit of the rapidity region $0.0 < |y| < 0.5$ and the p_T region 737 GeV to 2116 GeV. The upper left plot show a fit with the default scale(μ_r/μ_f) = (1.0/1.0) while the plot on the upper right side shows the fit with the maximal downward deviation on α_s observed with the scale variations (μ_r/μ_f) = (0.5/0.5). The plot in the midst shows the fit with the observed maximal upwards deviation with the scale variation (μ_r/μ_f) = (2.0/2.0).

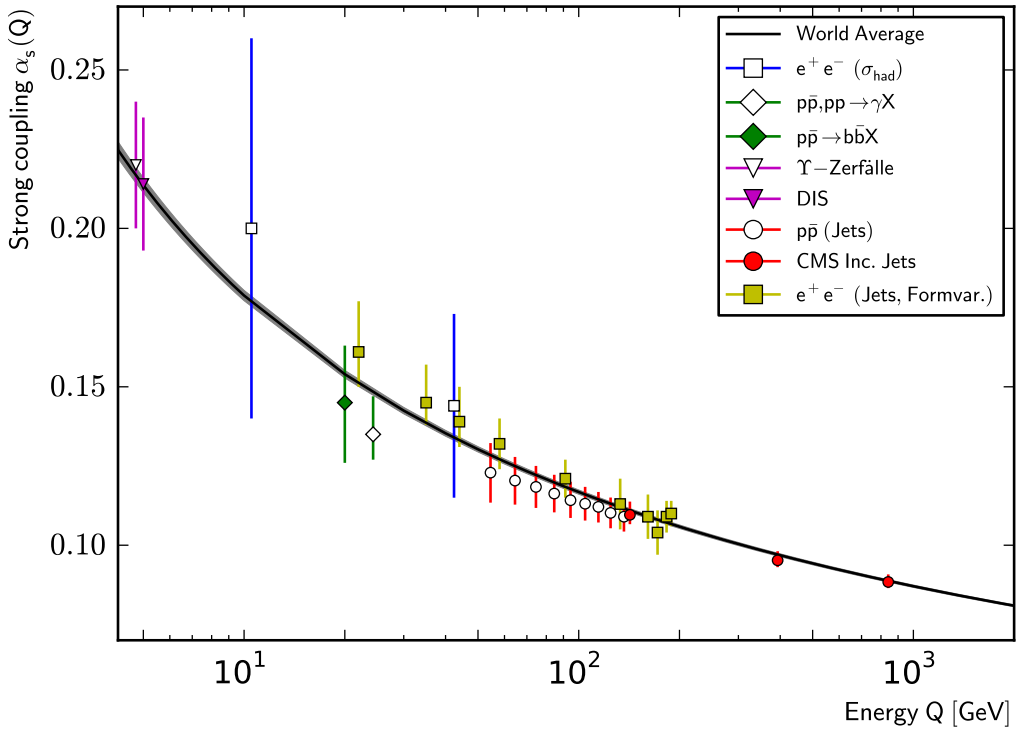


Figure 5.20.: Summary of various measurements of α_s evolved to the scale of the measurement. The results of the α_s extraction from inclusive jet cross sections are shown as well (red points). The α_s values shown with their overall uncertainty are in agreement with the predicted running of the strong coupling. The code for this plot is taken from [57] and uses data from [60, 61].

6. Conclusion and Outlook

Studies of proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ at the LHC opened a new energy range. QCD predictions have never been tested in this phase-space region and especially measurements of jet cross sections with high transverse momenta can probe QCD in these energy ranges. The CMS experiment recorded data of proton-proton collisions at the runs of the LHC in 2010 and 2011. The data sample recorded in 2011 corresponds to a an integrated luminosity of $4.67\,\text{fb}^{-1}$. The inclusive jet cross section has been measured differentially in the transverse momentum and the rapidity of the jets. This spectrum was unfolded and compared to NLO calculations which were found to be in agreement within uncertainties.

This thesis focused on the sensitivity of the PDFs and the strong coupling to the inclusive jet cross section. The recently released PDF fitting framework HERAFitter has been employed to study the impact of inclusive jet data on PDF fits.

The quark-PDFs are sensitive to the forward regions and the high-p_T central region of the inclusive jet measurement, while the gluon PDF is sensitive to the central rapidity region. The PDFs can be constrained further in the high-x phase-space region which is not accessible by HERA data. The uncertainties of the PDFs have been studied in detail and it was found that the inclusion of inclusive jet data in PDF fits significantly reduces the uncertainty in the high-x region.

The sensitivity of the strong coupling to the inclusive jet cross section provides the possibility to extract the strong coupling. To extract the strong coupling, PDF fits provided by the NNPDF collaboration has been used. These were determined for a series of fixed values of $\alpha_\text{s}(M_Z^2)$. The extracted values of $\alpha_\text{s}(M_Z^2)$ are in good agreement with similar measurements and also with the world average of $\alpha_\text{s}(M_Z^2)$.

Since the dataset covers a wide region in the transverse momentum of the jet, it was possible to extract the strong coupling in different energy regions. The extracted values of α_s are in agreement with the predicted running of α_s up to energies of 800 GeV.

The dominant experimental source of uncertainty results from the jet energy corrections. The ongoing studies of the jet energy scale and resolution corrections will further reduce the uncertainties on the experimental side. The dominant source on the theory side is the scale uncertainty due to limitations to NLO calculations. Calculations in NNLO are needed to reduce this source of uncertainty. Recent progress on the theory side raise the hope that these calculations will be available still this year.

This will lead to a large reduction of the uncertainty of the cross section and the extracted strong coupling. Furthermore, this will also favour the consideration of the strong coupling derived from jet cross sections in the determination of the world average of α_s.

A. Appendix

A.1. PDF Sensitivity Study

A.1.1. Technical Details of the Fit Procedure

The input parametrisation for the minimization using MINUIT is defined in steering cards. The starting parametrisation and the final parametrisation as determined by MINUIT is shown.

Fits with HERA-I DIS Data

The starting parametrisation and the minimisation strategy:

```
set title
new YOUR PDF
parameters
#  No. Parameter name                Value   Step size
   1   'Ag'                          0.0000   0.000000e+00
   2   'Bg'                         -0.13945  6.118400e-04
   3   'Cg'                           4.8548  2.730500e-02
   4   'Dg'                           0.0000   0.0
   5   'Eg'                           0.0000   0.0
   6   'Fg'                           0.0000   0.000000e+00
   7   'Aprig'                       0.16655   6.918900e-03
   8   'Bprig'                      -0.44303   1.866200e-03
   9   'Cprig'                        25.000   0.000000e+00
  11   'Auv'                          0.0000   0.000000e+00
  12   'Buv'                         0.64105   1.435500e-03
  13   'Cuv'                          5.4603   1.199800e-02
  14   'Duv'                          -1.7082   0.01
  15   'Euv'                          19.603   1.682400e-01
  21   'Adv'                          0.0000   0.000000e+00
  22   'Bdv'                         0.6184   6.207300e-03
  23   'Cdv'                          4.3605   4.840600e-02
  24   'Ddv'                          0.0000   0.0
  25   'Edv'                          0.0000   0.0
  31   'AUbar'                        0.0000   0.000000e+00
  32   'BUbar'                        0.0000   0.000000e+00
```

```
  33   'CUbar'                    1.5898   4.800100e-03
  34   'DUbar'                    0.0000   0.0
  35   'EUbar'                    0.0000   0.0
  41   'ADbar'                   0.1550   3.959200e-04
  42   'BDbar'                  -0.1695   3.944400e-04
  43   'CDbar'                    3.6181   6.578000e-02
  44   'DDbar'                    0.0000   0.0
  45   'EDbar'                    0.0000   0.0
 101   'alphas'                  0.11760   0.000000e+00
 102   'fs'                      0.31000   0.000000e+00
 103   'fcharm'                   0.0000   0.000000e+00

simplex 10000
migrad 200000
hesse
return
```

The final parametrisation of the successful minimisation of a fit to HERA-I DIS data.

```
 **********
 **    3 **HESSE
 **********
 COVARIANCE MATRIX CALCULATED SUCCESSFULLY

 FCN=   566.1116 FROM HESSE STATUS=OK 148 CALLS  4765 TOTAL
                 EDM=  0.82E-05 STRATEGY= 1 ERROR MATRIX ACCURATE

  EXT PARAMETER                                   INTERNAL      INTERNAL
  NO.   NAME        VALUE          ERROR        STEP SIZE       VALUE
   1    Ag          0.0000       constant
   2    Bg        -0.10050       0.18808       0.31601E-05  -0.10050
   3    Cg          9.3734        1.8878       0.72688E-04    9.3734
   4    Dg          0.0000       constant
   5    Eg          0.0000       constant
   6    Fg          0.0000       constant
   7    Aprig      0.96415        1.2261       0.18643E-04   0.96415
   8    Bprig     -0.25844       0.66425E-01   0.31290E-05  -0.25844
   9    Cprig       25.000       constant
  11    Auv         0.0000       constant
  12    Buv        0.79840       0.61658E-01   0.41993E-05   0.79840
  13    Cuv         5.8996       0.51081       0.27580E-04    5.8996
  14    Duv        -3.1392       0.93897       0.10413E-03   -3.1392
  15    Euv         18.451        4.1739       0.24603E-03    18.451
```

```
 21    Adv          0.0000       constant
 22    Bdv         0.73473       0.10517       0.14744E-04   0.73473
 23    Cdv          6.1250        1.1307       0.14123E-03    6.1250
 24    Ddv          0.0000       constant
 25    Edv          0.0000       constant
 31    AUbar        0.0000       constant
 32    BUbar        0.0000       constant
 33    CUbar        2.2148       0.57922       0.73341E-04    2.2148
 34    DUbar        0.0000       constant
 35    EUbar        0.0000       constant
 41    ADbar       0.15085       0.84168E-02   0.42413E-06   0.15085
 42    BDbar      -0.17481       0.73584E-02   0.13156E-05  -0.17481
 43    CDbar        1.1354       0.31203       0.23665E-04    1.1354
 44    DDbar        0.0000       constant
 45    EDbar        0.0000       constant
101    alphas      0.11760       constant
102    fs          0.31000       constant
103    fcharm       0.0000       constant
```

Fits with HERA-I DIS + CMS Jets 2011 Data.

The starting parametrisation and the minimisation strategy:

```
set title
new YOUR PDF
parameters
#  No. Parameter name              Value    Step size
   1   'Ag'                         0.0000   0.000000e+00
   2   'Bg'                        -0.13945  6.118400e-04
   3   'Cg'                         4.8548   2.730500e-02
   4   'Dg'                         0.0000   0.0
   5   'Eg'                         0.0000   0.0
   6   'Fg'                         0.0000   0.000000e+00
   7   'Aprig'                      0.16655  6.918900e-03
   8   'Bprig'                     -0.44303  1.866200e-03
   9   'Cprig'                      25.000   0.000000e+00
  11   'Auv'                        0.0000   0.000000e+00
  12   'Buv'                        0.64105  1.435500e-03
  13   'Cuv'                        5.4603   1.199800e-02
  14   'Duv'                       -1.7082    0.01
  15   'Euv'                        19.603   1.682400e-01
  21   'Adv'                        0.0000   0.000000e+00
  22   'Bdv'                        0.6184   6.207300e-03
  23   'Cdv'                        4.3605   4.840600e-02
```

```
  24   'Ddv'                      0.0000   0.0
  25   'Edv'                      0.0000   0.0
  31   'AUbar'                    0.0000   0.000000e+00
  32   'BUbar'                    0.0000   0.000000e+00
  33   'CUbar'                    1.5898   4.800100e-03
  34   'DUbar'                    0.0000   0.0
  35   'EUbar'                    0.0000   0.0
  41   'ADbar'                   0.1550   3.959200e-04
  42   'BDbar'                  -0.1695   3.944400e-04
  43   'CDbar'                    3.6181   6.578000e-02
  44   'DDbar'                    0.0000   0.0
  45   'EDbar'                    0.0000   0.0
 101   'alphas'                  0.11760   0.000000e+00
 102   'fs'                      0.31000   0.000000e+00
 103   'fcharm'                   0.0000   0.000000e+00

simplex 10000
migrad 200000
hesse
return
```

The final parametrisation of the successful minimisation of a fit to HERA-I DIS and CMS inclusive jet data.

```
**********
 **    3 **HESSE
 **********
 COVARIANCE MATRIX CALCULATED SUCCESSFULLY

 FCN=   756.0989 FROM HESSE STATUS=OK 150 CALLS 1749 TOTAL
                 EDM=  0.63E-06 STRATEGY= 1 ERROR MATRIX ACCURATE

  EXT PARAMETER                                   INTERNAL      INTERNAL
  NO.   NAME        VALUE          ERROR        STEP SIZE       VALUE
   1    Ag          0.0000       constant
   2    Bg        -0.13219       0.11248       0.20945E-04  -0.13219
   3    Cg          5.0016       0.50864       0.30821E-03    5.0016
   4    Dg          0.0000       constant
   5    Eg          0.0000       constant
   6    Fg          0.0000       constant
   7    Aprig      0.13618       0.16805       0.26403E-04   0.13618
   8    Bprig     -0.46106       0.80772E-01   0.33103E-04  -0.46106
   9    Cprig       25.000       constant
  11    Auv         0.0000       constant
```

12	Buv	0.63921	0.26053E-01	0.26057E-04	0.63921
13	Cuv	5.5431	0.18223	0.18601E-03	5.5431
14	Duv	-1.9465	0.63275	0.91114E-03	-1.9465
15	Euv	20.723	2.0345	0.21608E-02	20.723
21	Adv	0.0000	constant		
22	Bdv	0.61230	0.77768E-01	0.88762E-04	0.61230
23	Cdv	4.7840	0.68149	0.80207E-03	4.7840
24	Ddv	0.0000	constant		
25	Edv	0.0000	constant		
31	AUbar	0.0000	constant		
32	BUbar	0.0000	constant		
33	CUbar	1.6203	0.77644E-01	0.31644E-03	1.6203
34	DUbar	0.0000	constant		
35	EUbar	0.0000	constant		
41	ADbar	0.15663	0.41838E-02	0.10187E-04	0.15663
42	BDbar	-0.16756	0.42676E-02	0.70124E-05	-0.16756
43	CDbar	2.7119	0.51118	0.51042E-03	2.7119
44	DDbar	0.0000	constant		
45	EDbar	0.0000	constant		
101	alphas	0.11760	constant		
102	fs	0.31000	constant		
103	fcharm	0.0000	constant		

A.1.2. PDF Distributions

Fits with inclusive jet data 2010 from CMS

The PDFs of the light anti-quark PDFs are shown in the Figures A.1 and A.2. Both PDFs show no big impact by the inclusive jet data. This results from the suppressed $q\bar{q}$-subprocess in the jet production.

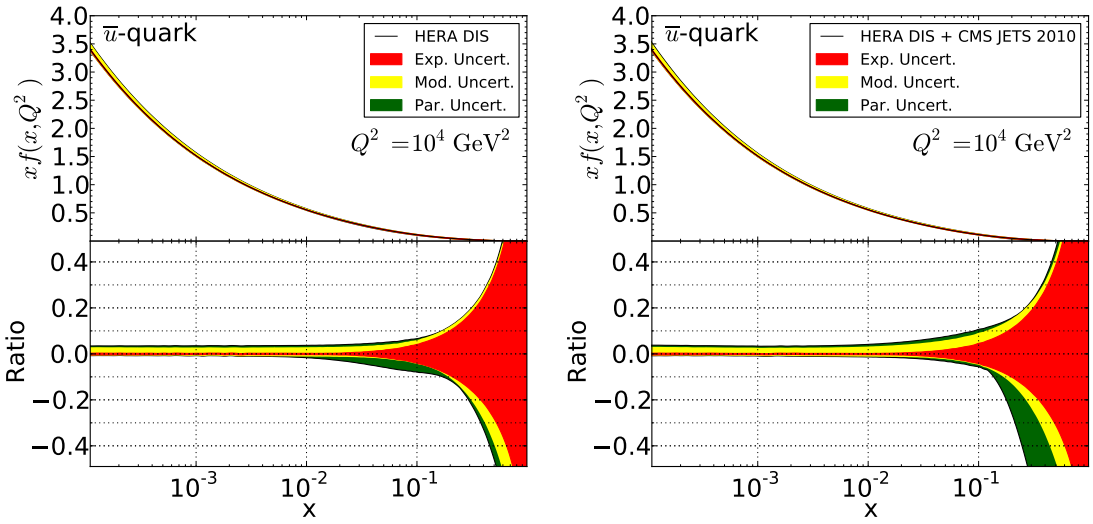


Figure A.1.: The anti up-quark PDF. The q-$\bar{q}$ subprocess is suppressed and the impact on the PDFs is negligible. The PDF is evolved to $Q^2 = 10^4\,\mathrm{GeV}^2$.

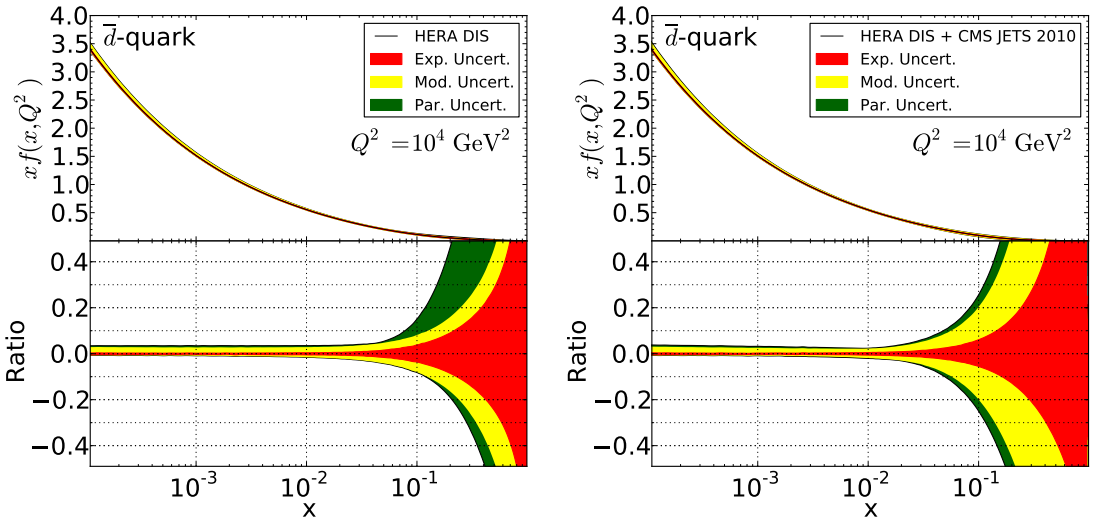


Figure A.2.: The anti down-quark PDF. The $q\bar{q}$ subprocess is suppressed and the impact on the PDFs is negligible. The PDF is evolved to $Q^2 = 10^4\,\mathrm{GeV}^2$.

Fits with inclusive jet data 2011 from CMS

Similar to the 2010 data fits, the suppressed $q\bar{q}$-subprocess in the jet production should result in a negligible impact on the anti-quark PDFs, While this remains true for the anti down-quark PDF with negligible effect, there is a non-negligible impact on the anti up-quark PDF. The result is not fully understood, but it is possible that constrains of the QCD sum rules lead to the reduction of the uncertainty visible in the high-x region.

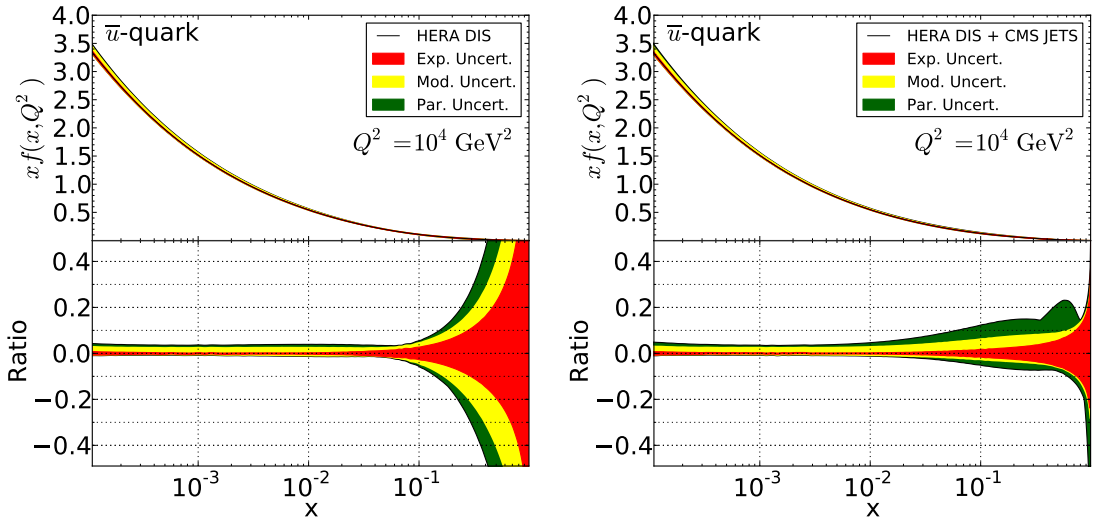


Figure A.3.: The anti up-quark PDF. The $q\bar{q}$ processes in the inclusive jet production are suppressed. Nonetheless, a reduction of the model and experimental uncertainties is visible. The PDF is evolved to $Q^2 = 10^4\,\mathrm{GeV}^2$.

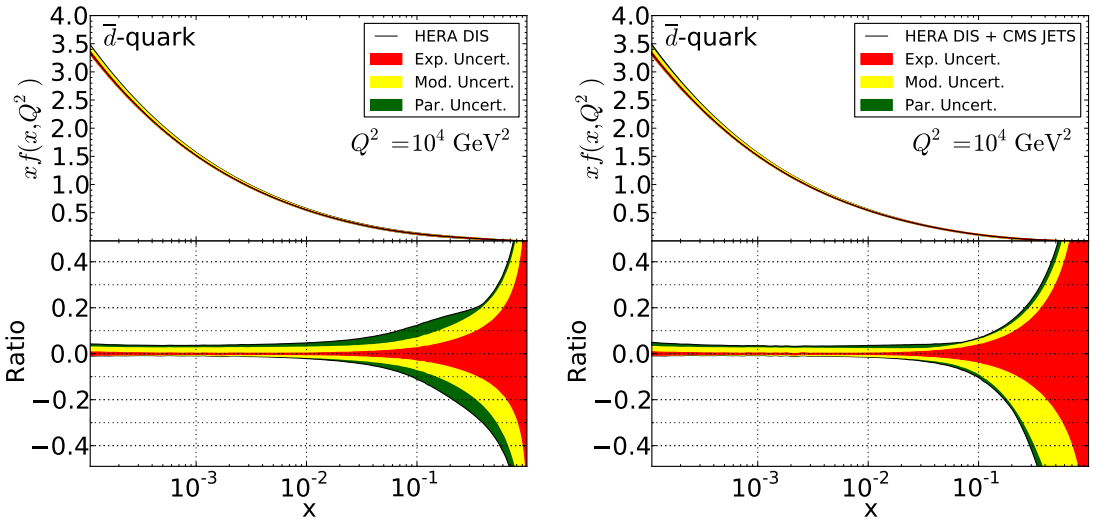


Figure A.4.: The anti down-quark PDF. The $q\bar{q}$ processes in the inclusive jet production are suppressed and the impact on the PDF is negligible. The PDF is evolved to $Q^2 = 10^4\,\mathrm{GeV}^2$.

A.1.3. Ratios

Fits with inclusive jet data 2010 from CMS

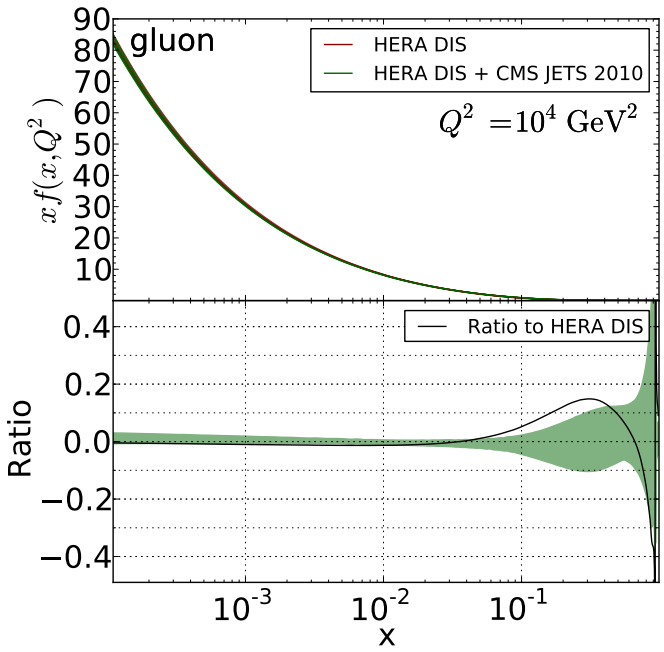


Figure A.5.: Ratio of the gluon PDF in a fit with and without CMS 2010 inclusive jet data. The lower plot shows the total uncertainty calculated by adding experimental, model and parametrization uncertainties in quadrature.

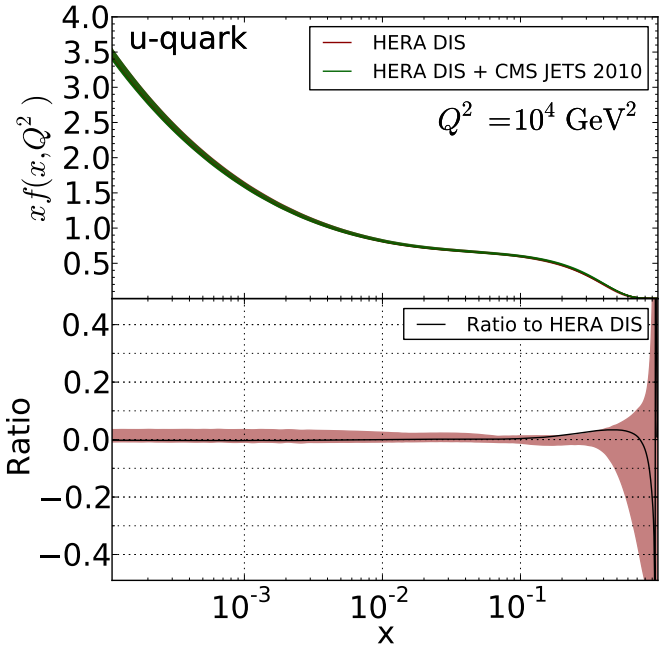


Figure A.6.: Ratio of the up-quark PDF in a fit with and without CMS 2010 inclusive jet data. The lower plot shows the total uncertainty calculated by adding experimental, model and parametrization uncertainties in quadrature.

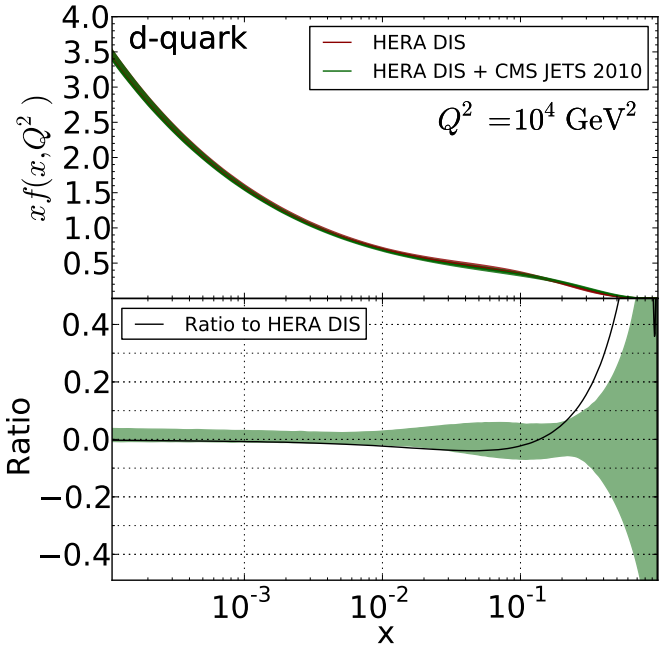


Figure A.7.: Ratio of the down-quark PDF in a fit with and without 2010 inclusive jet data. The lower plot shows the total uncertainty calculated by adding experimental, model and parametrization uncertainties in quadrature.

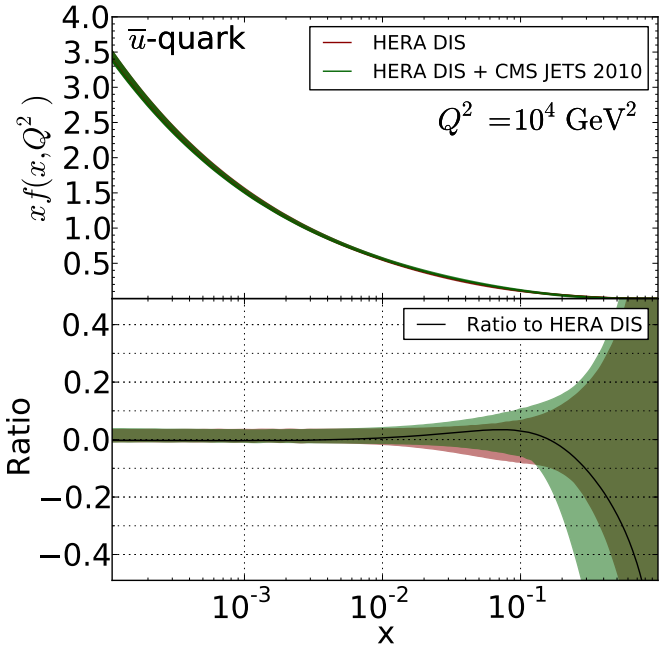


Figure A.8.: Ratio of the $\bar{u}$-quark PDF in a fit with and without 2010 inclusive jet data. The lower plot shows the total uncertainty calculated by adding experimental, model and parametrization uncertainties in quadrature.

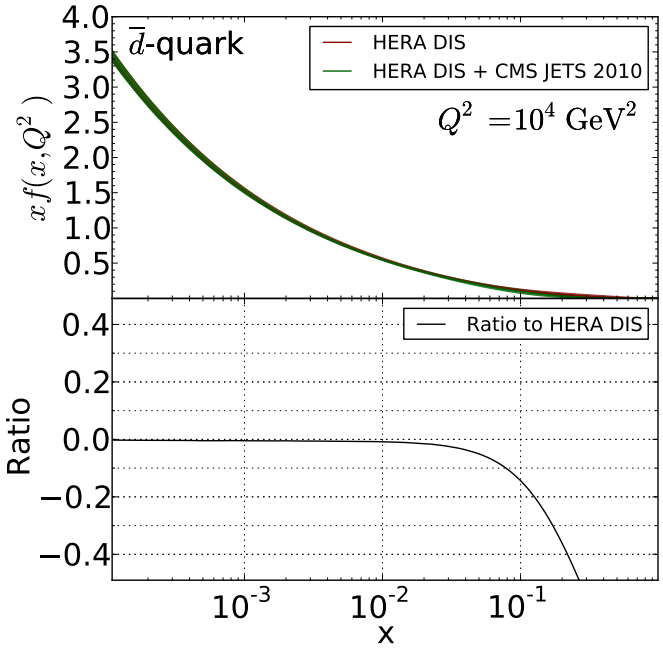


Figure A.9.: Ratio of the $\bar{d}$-quark PDF in a fit with and without 2010 inclusive jet data. The lower plot shows the total uncertainty calculated by adding experimental, model and parametrization uncertainties in quadrature.

Fits with inclusive jet data 2011 from CMS

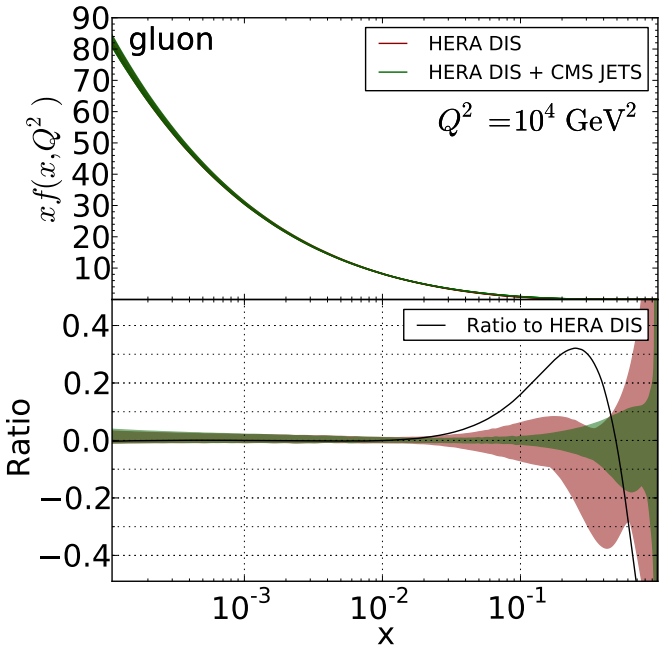


Figure A.10.: Ratio of the gluon PDF in a fit with and without 2011 inclusive jet data. The lower plot shows the total uncertainty calculated by adding experimental, model and parametrization uncertainties in quadrature.

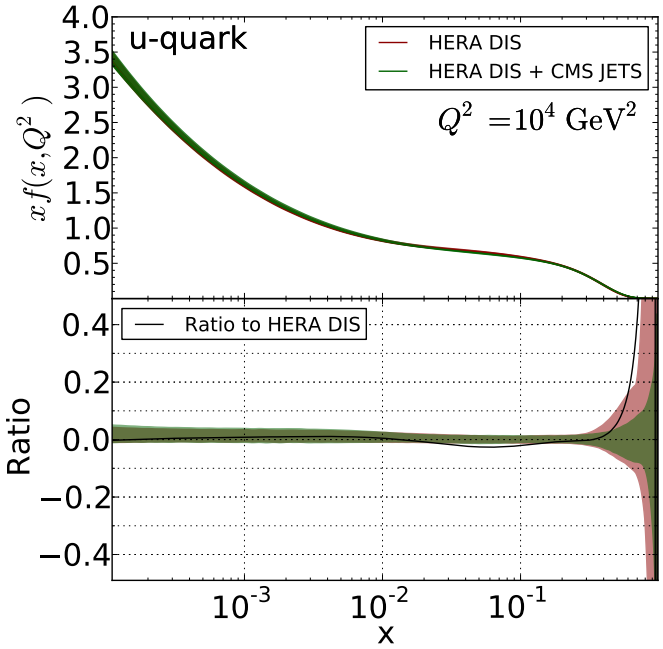


Figure A.11.: Ratio of the up-quark PDF in a fit with and without 2011 inclusive jet data. The lower plot shows the total uncertainty calculated by adding experimental, model and parametrization uncertainties in quadrature.

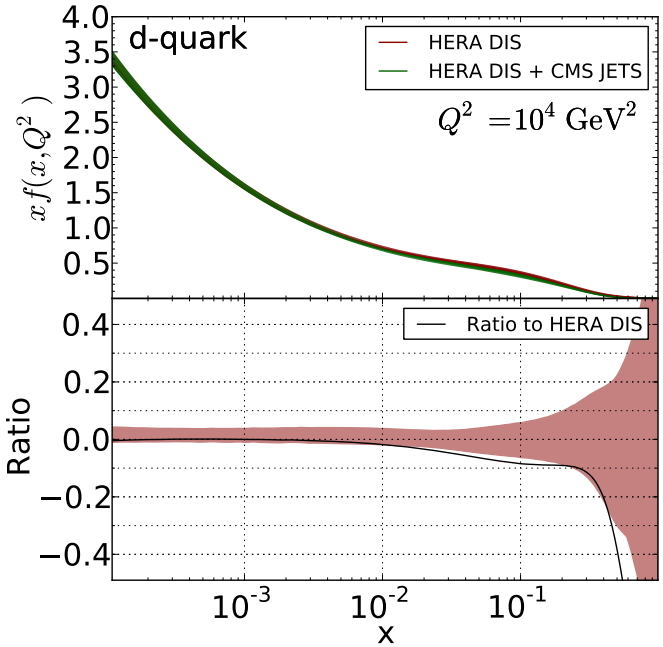


Figure A.12.: Ratio of the down-quark PDF in a fit with and without 2011 inclusive jet data. The lower plot shows the total uncertainty calculated by adding experimental, model and parametrization uncertainties in quadrature.

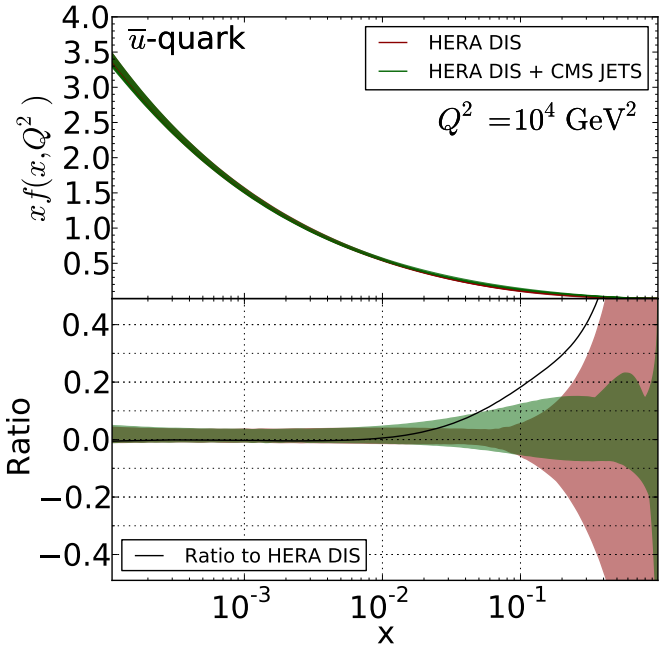


Figure A.13.: Ratio of the $\bar{u}$-quark PDF in a fit with and without 2011 inclusive jet data. The lower plot shows the total uncertainty calculated by adding experimental, model and parametrization uncertainties in quadrature.

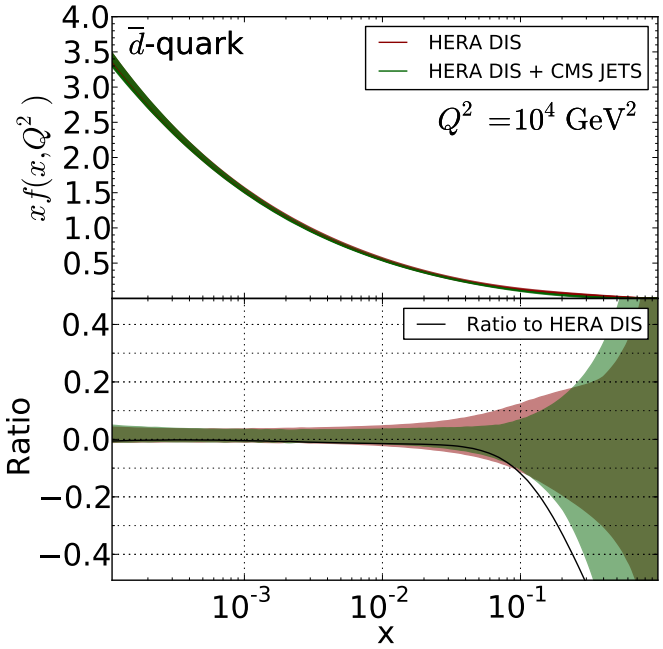


Figure A.14.: Ratio of the $\bar{d}$-quark PDF in a fit with and without 2011 inclusive jet data. The lower plot shows the total uncertainty calculated by adding experimental, model and parametrization uncertainties in quadrature.

A.2. Extraction of α_s

A.2.1. Sensitivity of α_s

The sensitivity on α_s by various PDF sets is shown in this chapter. Predictions with the global PDF sets NNPDF21, CT10 and MSTW2008 describe the data quite well and an extraction of α_s is possible. The PDF set ABM11 does not describe the data well, especially in the high-p_T region. This behaviour of the ABM11 PDF Set is known and also observed in other analyses.

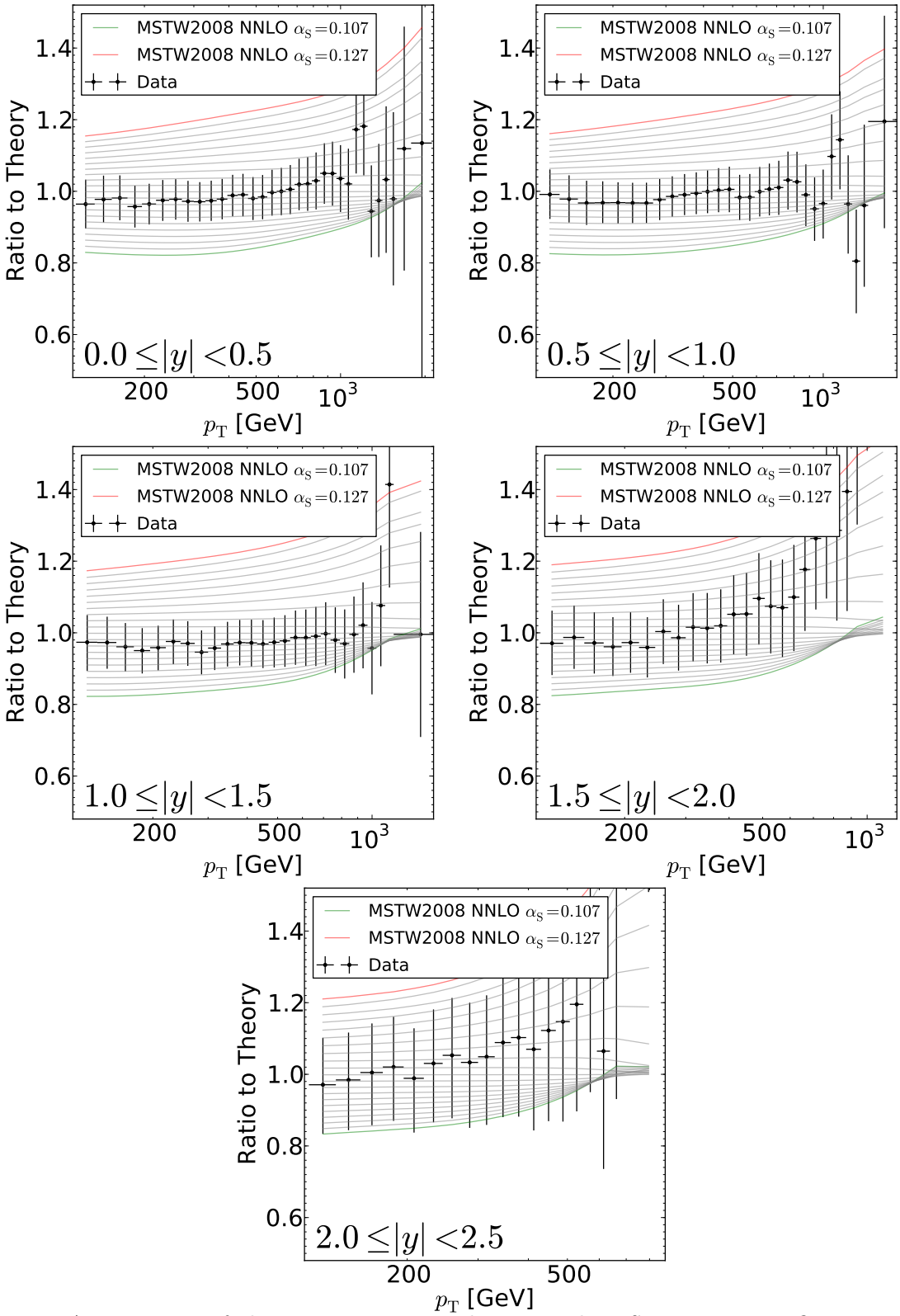


Figure A.15.: Ratio of the cross section prediction with *MSTW2008 NNLO* using all α_s-variation fits to the standard fit. The fits for fixed values of α_s are available in a range of 0.107 to 0.127 in steps of 0.001. The data points with the total experimental uncertainties are shown as well.

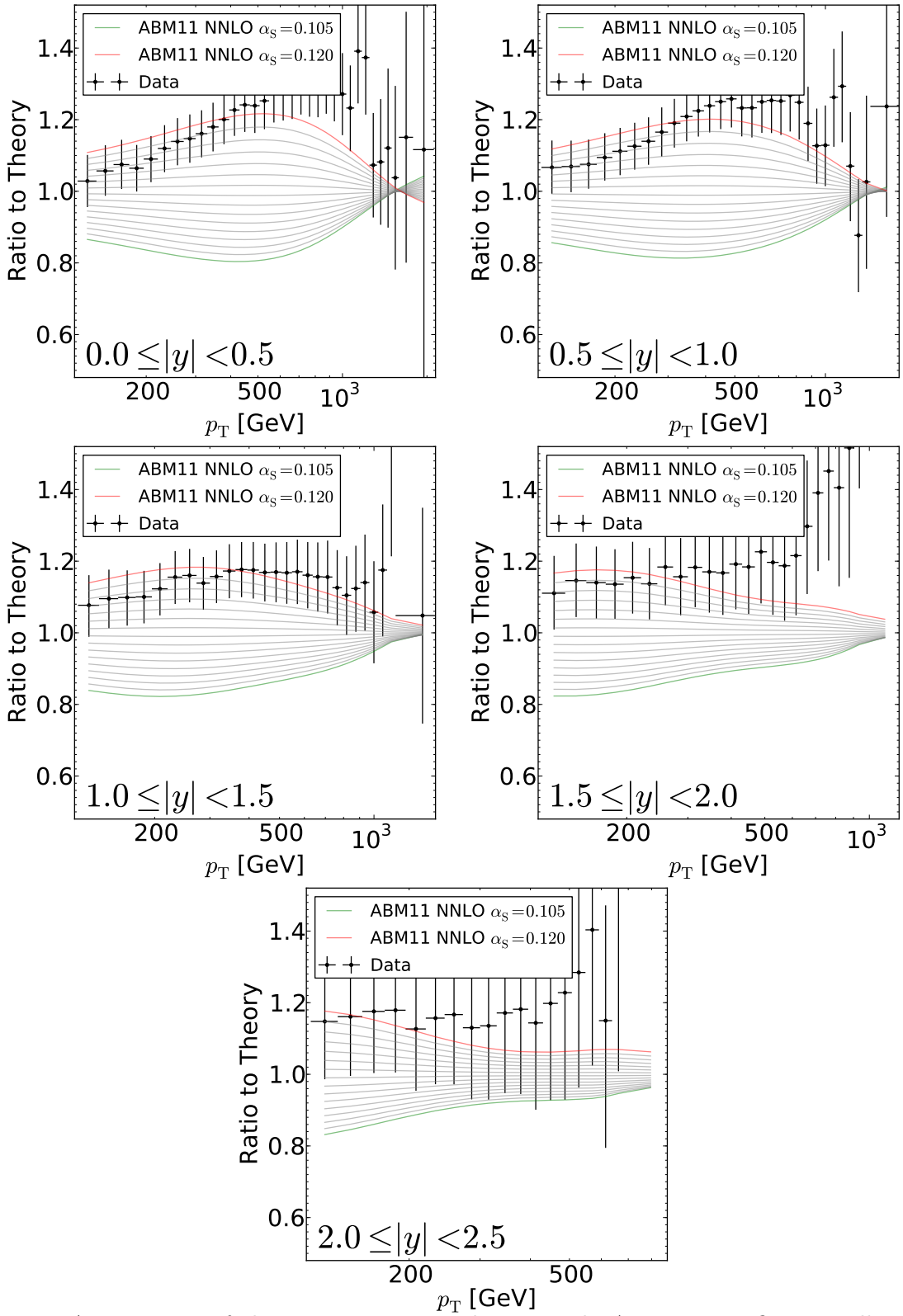


Figure A.16.: Ratio of the cross section prediction with *ABM11 NNLO* using all α_s-variation fits to the standard fit. The fits for fixed values of α_s are available in a range of 0.105 to 0.120 in steps of 0.001. The data points with the total experimental uncertainties are shown as well.

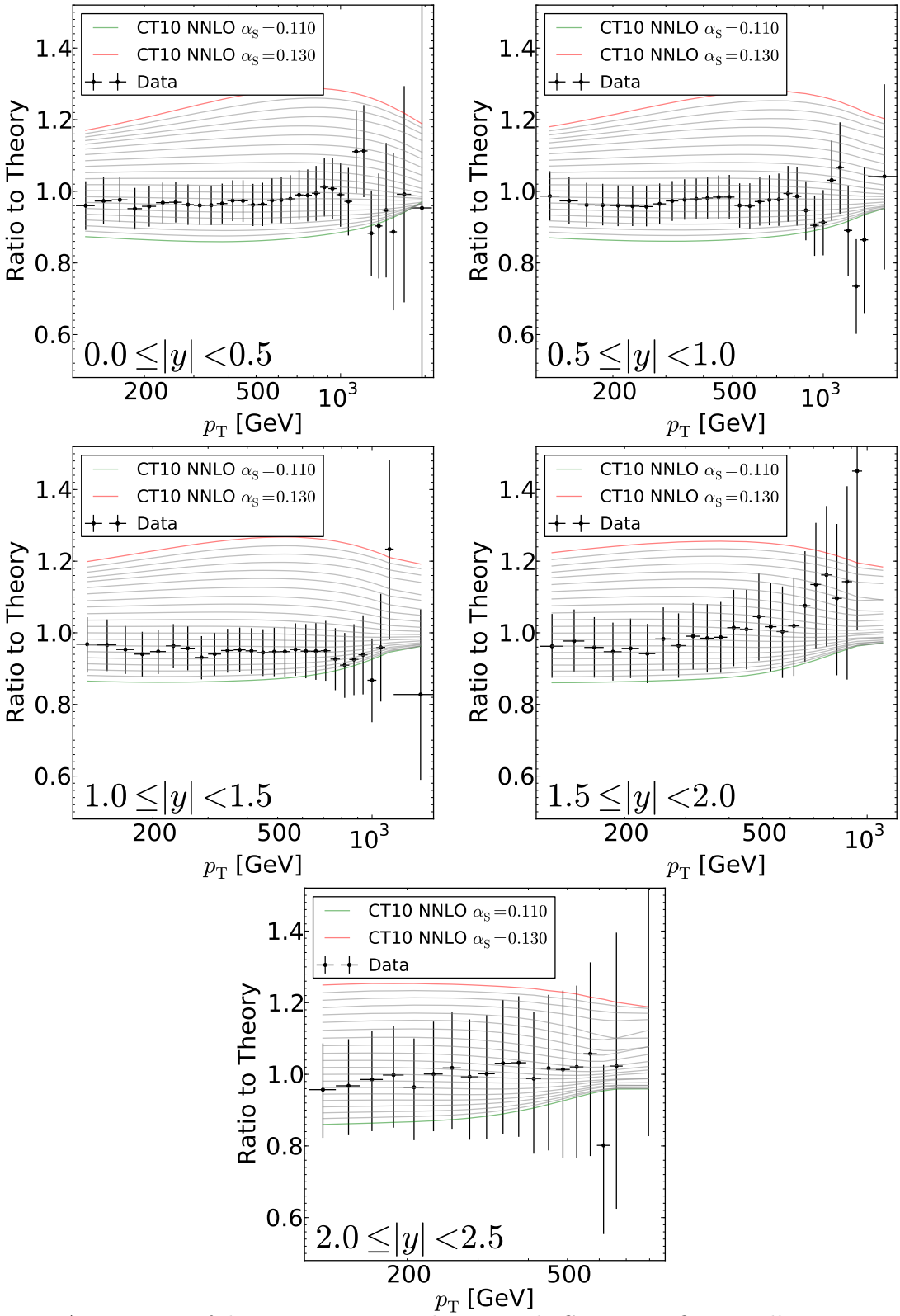


Figure A.17.: Ratio of the cross section prediction with *CT10 NNLO* using all α_s-variation fits to the standard fit. The fits for fixed values of α_s are available in a range of 0.110 to 0.130 in steps of 0.001. The data points with the total experimental uncertainties are shown as well.

A.2.2. 6p Scale-Variation

The scale uncertainty is calculated using independent scale factors variations of 0.5 and 2.0 for the scales μ_r and μ_f. All considered six scale variations are shown for the five rapidity bins.

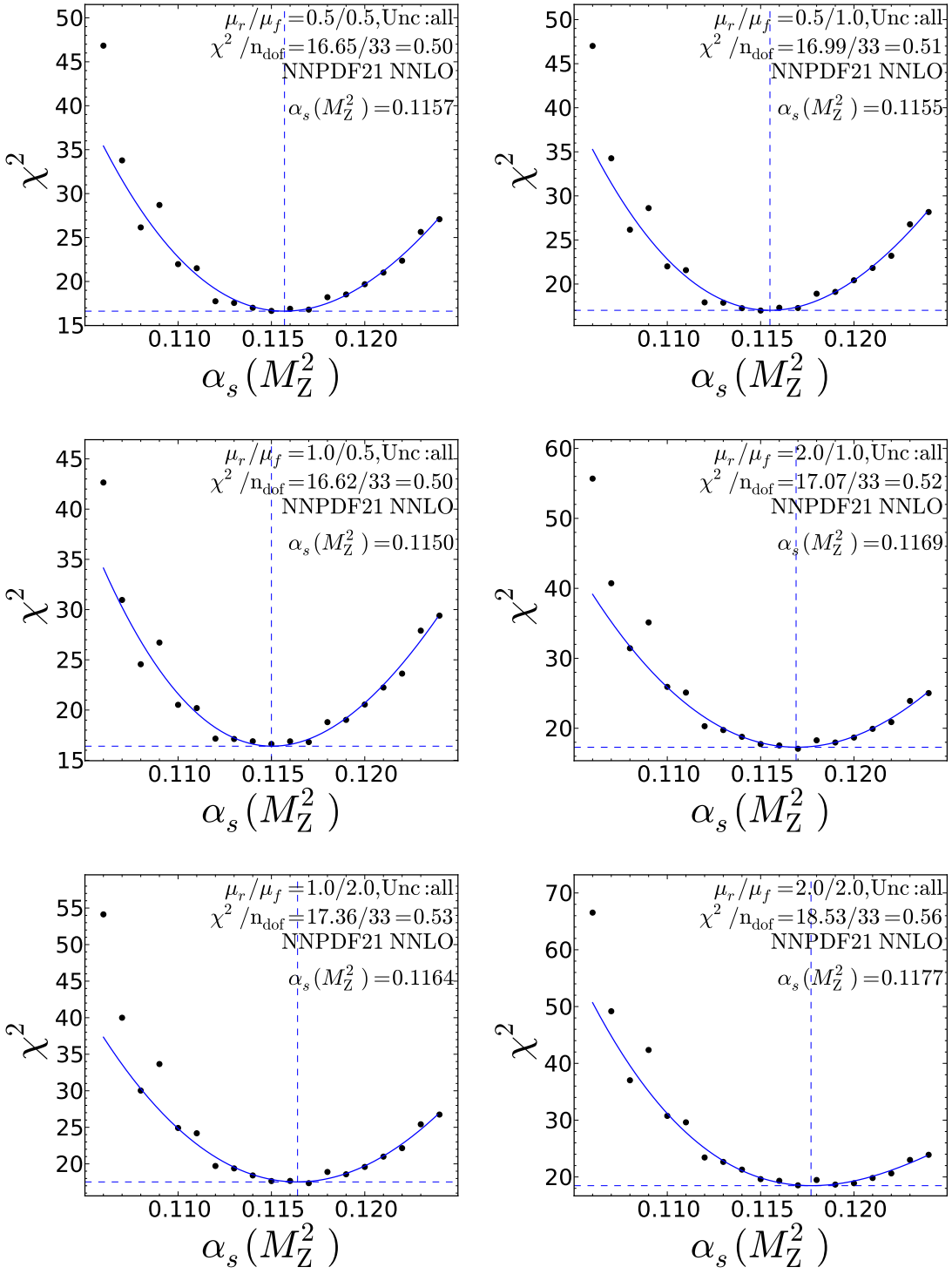


Figure A.18.: α_s fit of the rapdity region rapidity region $\mathbf{0.0 < |y| < 0.5}$ with the NNPDF 2.1 NNLO set. The figures show the six independent scale variations and the impact on the extracted α_s value. The maximum and minimum deviation of the α_s value define the scale uncertainty.

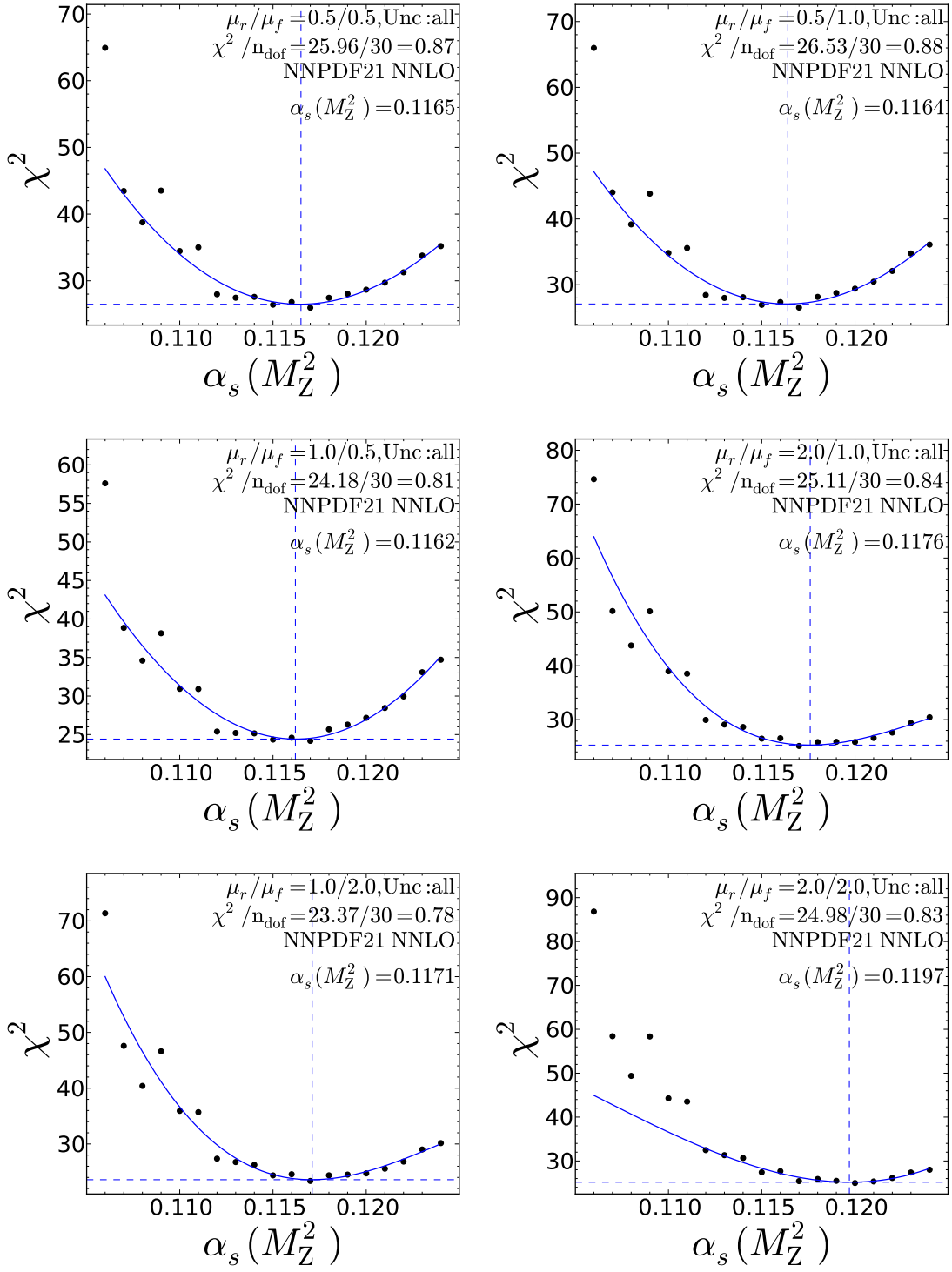


Figure A.19.: α_s fit of the rapdity region rapidity region $\mathbf{0.5 < |y| < 1.0}$ with the NNPDF 2.1 NNLO set. The figures show the six independent scale variations and the impact on the extracted α_s value. The maximum and minimum deviation of the α_s value define the scale uncertainty.

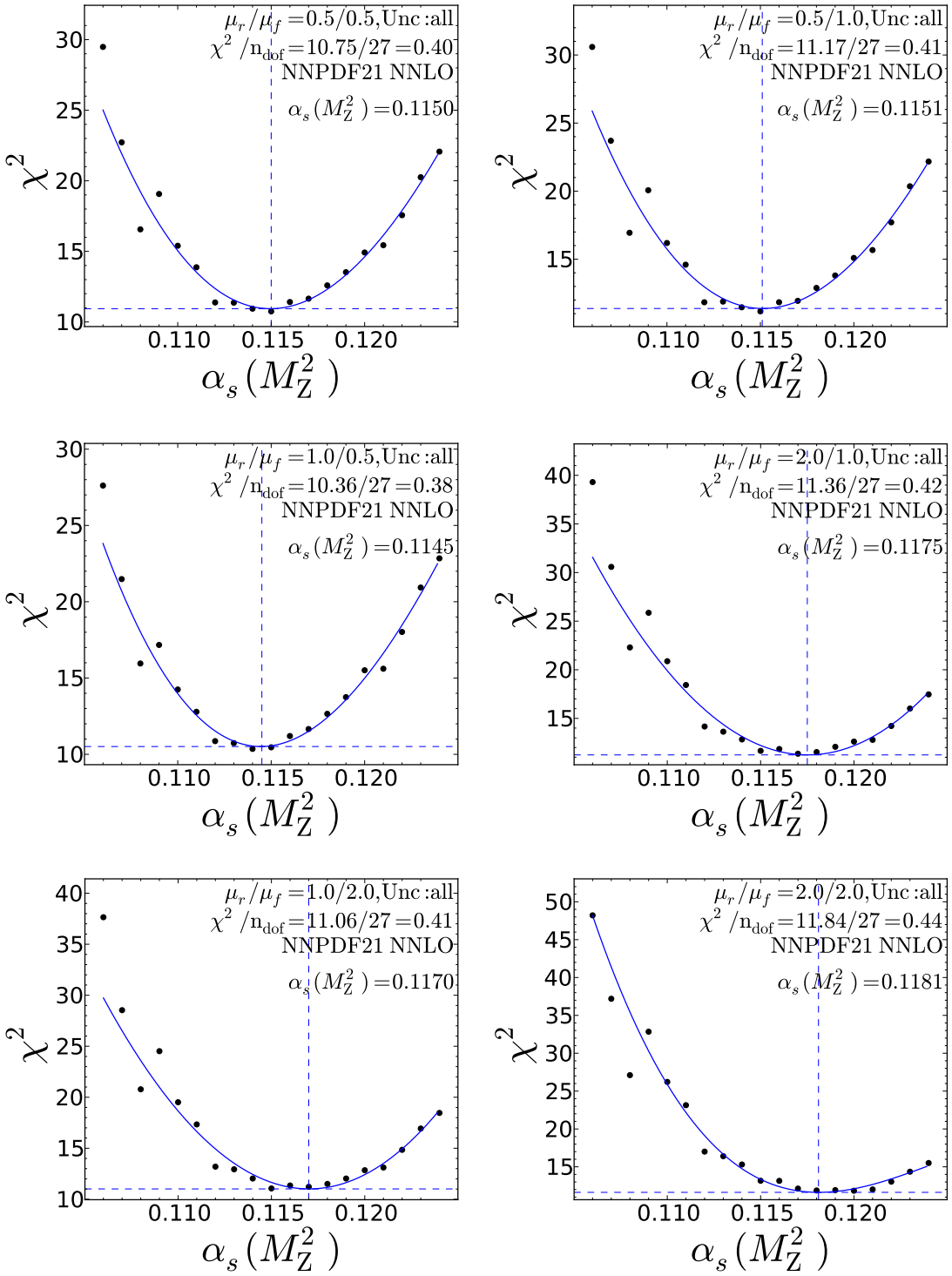


Figure A.20.: α_s fit of the rapdity region rapidity region $\mathbf{1.0 < |y| < 1.5}$ with the NNPDF 2.1 NNLO set. The figures show the six independent scale variations and the impact on the extracted α_s value. The maximum and minimum deviation of the α_s value define the scale uncertainty.

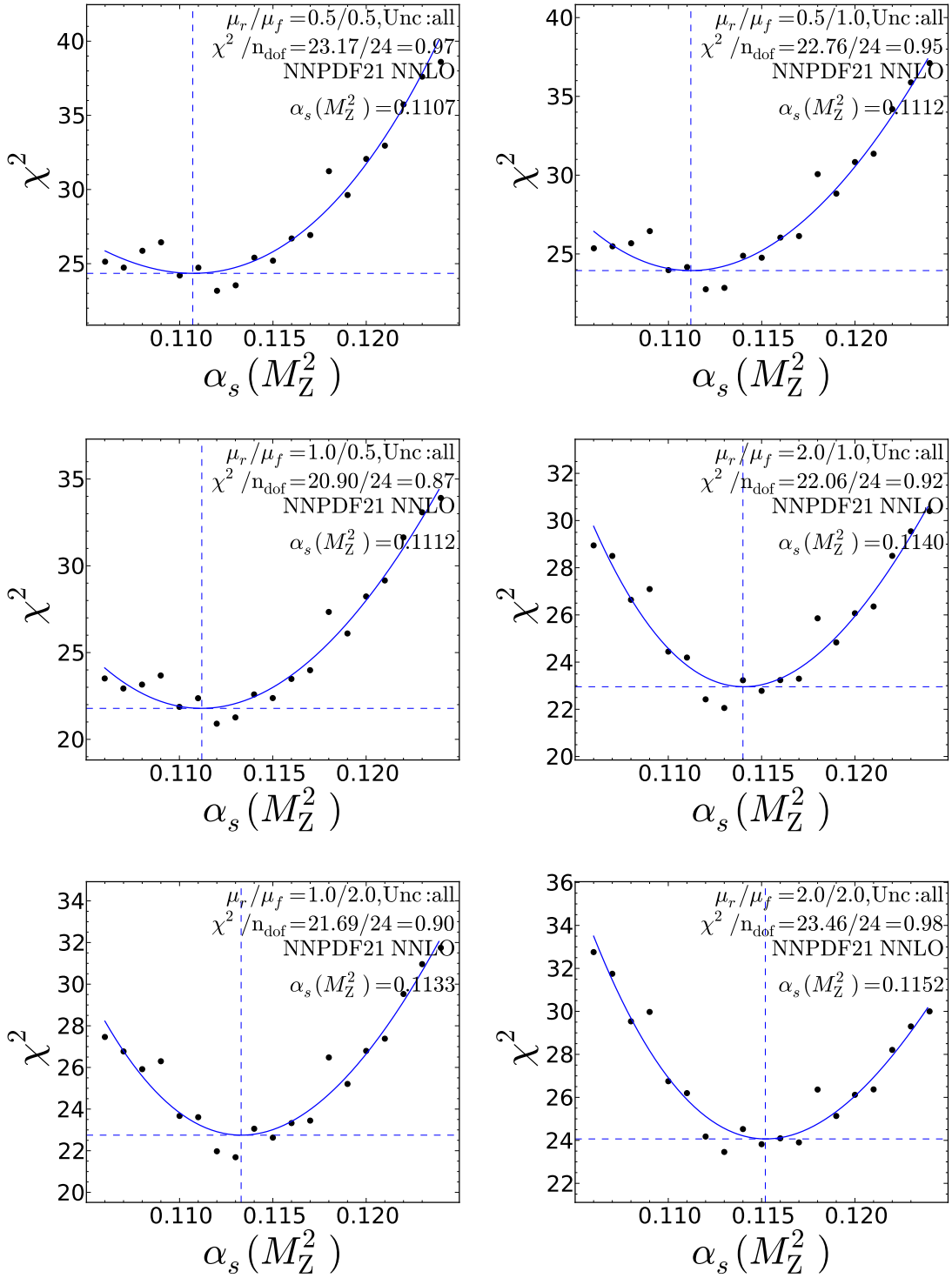


Figure A.21.: α_s fit of the rapdity region rapidity region $\mathbf{1.5 < |y| < 2.0}$ with the NNPDF 2.1 NNLO set. The figures show the six independent scale variations and the impact on the extracted α_s value. The maximum and minimum deviation of the α_s value define the scale uncertainty.

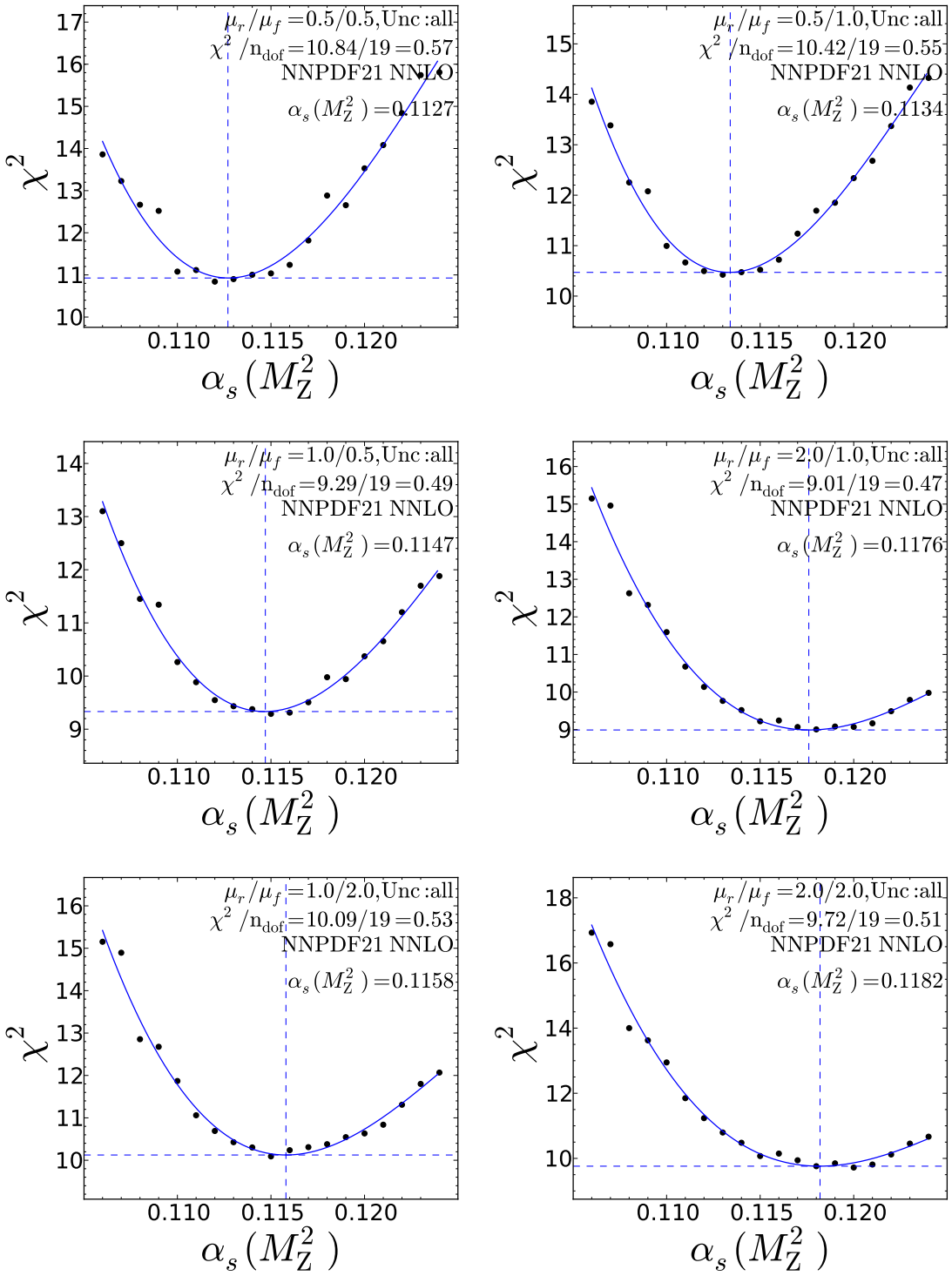


Figure A.22.: α_s fit of the rapdity region rapidity region $\mathbf{2.0 < |y| < 2.5}$ with the NNPDF 2.1 NNLO set. The figures show the six independent scale variations and the impact on the extracted α_s value. The maximum and minimum deviation of the α_s value define the scale uncertainty.

List of Figures

List of Tables

Bibliography

[1] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys.Lett.*, B716:1–29, 2012.

[2] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012.

[3] Steven Weinberg. Nonabelian Gauge Theories of the Strong Interactions. *Phys.Rev.Lett.*, 31:494–497, 1973.

[4] H. Fritzsch, Murray Gell-Mann, and H. Leutwyler. Advantages of the Color Octet Gluon Picture. *Phys.Lett.*, B47:365–368, 1973.

[5] D.J. Gross and Frank Wilczek. Asymptotically Free Gauge Theories. 1. *Phys.Rev.*, D8:3633–3652, 1973.

[6] William A. Bardeen, A. J. Buras, D. W. Duke, and T. Muta. Deep-inelastic scattering beyond the leading order in asymptotically free gauge theories. *Phys. Rev. D*, 18:3998–4017, Dec 1978.

[7] J. Beringer et al. Review of Particle Physics (RPP). *Phys.Rev.*, D86:010001, 2012.

[8] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov.Phys.JETP*, 46:641–653, 1977.

[9] V.N. Gribov and L.N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov.J.Nucl.Phys.*, 15:438–450, 1972.

[10] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl.Phys.*, B126:298, 1977.

[11] Sasha Glazov. Measurement of DIS cross section at HERA. *Braz.J.Phys.*, 37:793–797, 2007.

[12] A.D. Martin, W.J. Stirling, R.S. Thorne, and G. Watt. Parton distributions for the lhc. *Eur.Phys.J.*, C63:189–285, 2009.

[13] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, et al. New parton distributions for collider physics. *Phys.Rev.*, D82:074024, 2010.

[14] Richard D. Ball, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Jose I. Latorre, et al. A first unbiased global NLO determination of parton distributions and their uncertainties. *Nucl.Phys.*, B838:136–206, 2010.

[15] S. Alekhin, J. Blumlein, and S. Moch. PDF fit in the fixed-flavor-number scheme. *Nucl.Phys.Proc.Suppl.*, 222-224:41–51, 2012.

[16] F.D. Aaron et al. Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA. *JHEP*, 1001:109, 2010.

[17] Joram Berger. *Analysis of Z to mumu + Jet Events at the LHC*. PhD thesis, University of Karlsruhe, 2010. IEKP-KA/2010-23.

[18] Lyndon Evans et al. LHC Machine. *JINST*, 3:S08001, 2008.

[19] Schematic view of the LHC accelerator Complex. `http://public.web.cern.ch/public/en/Research/AccelComplex-en.html`. Accessed: Jan 2013.

[20] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008.

[21] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[22] A. Augusto Alves et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.

[23] The TOTEM Collaboration. The TOTEM Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08007, 2008.

[24] Overview of the CMS detector design. `http://cms.web.cern.ch/news/cms-detector-design`. Accessed: Jan 2013.

[25] Slice of the CMS detector. `https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/ShowDocument?docid=5697`. Accessed: Jan 2013.

[26] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[27] G. L. Bayatian and Chatrchyan. CMS Physics Technical Design Report Volume I: Detector Performance and Software. 2006.

[28] A. Scheurer. German contributions to the CMS computing infrastructure. *J.Phys.Conf.Ser.*, 219:062064, 2010.

[29] Root analysis framework. `http://root.cern.ch/drupal/content/about`. Accessed: Jan 2013.

[30] I. Antcheva and M. Ballintijn. ROOT, A C++ framework for petabyte data storage, statistical analysis and visualization. *Computer Physics Communications*, 2009.

[31] Grid-Control - An Alternative Grid Job Submission Tool, `https://ekptrac.physik.uni-karlsruhe.de/trac/grid-control/`.

[32] F.D. Aaron et al. Measurement of the Inclusive ep Scattering Cross Section at Low Q2 and x at HERA. *Eur.Phys.J.*, C63:625–678, 2009.

[33] S. Chekanov et al. A ZEUS next-to-leading-order QCD analysis of data on deep inelastic scattering. *Phys.Rev.*, D67:012007, 2003.

[34] C. Adloff et al. Measurement and QCD analysis of neutral and charged current cross-sections at HERA. *Eur.Phys.J.*, C30:1–32, 2003.

[35] Krzysztof Nowak. HERAFitter theory interfaces, `http://indico.cern.ch/getFile.py/access?contribId=5&resId=0&materialId=slides&confId=168468`.

[36] M.R. Whalley, D. Bourilkov, and R.C. Group. The Les Houches accord PDFs (LHAPDF) and LHAGLUE. 2005.

[37] M. Botje. QCDNUM: Fast QCD Evolution and Convolution. *Comput.Phys.Commun.*, 182:490–532, 2011.

[38] Zoltán Nagy. Three-jet cross sections in hadron-hadron collisions at next-to-leading order. *Phys. Rev. Lett.*, 88:122003, Mar 2002.

[39] T. Kluge, K. Rabbertz, and M. Wobisch. FastNLO: Fast pQCD calculations for PDF fits. pages 483–486, 2006.

[40] Z. Sullivan and Pavel M. Nadolsky. Heavy quark parton distribution functions and their uncertainties. *eConf*, C010630:P511, 2001.

[41] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, et al. Uncertainties of predictions from parton distribution functions. 2. The Hessian method. *Phys.Rev.*, D65:014013, 2001.

[42] G. D'Agostini. Bayesian reasoning in data analysis: A critical introduction. 2003.

[43] D. Stump, J. Pumplin, R. Brock, D. Casey, J. Huston, et al. Uncertainties of predictions from parton distribution functions. 1. The Lagrange multiplier method. *Phys.Rev.*, D65:014012, 2001.

[44] Daniel Stump. Uncertainties of Parton Distribution Functions. *eConf*, C030908:THAT005, 2003.

[45] M. Botje. Error estimates on parton density distributions. *J.Phys.*, G28:779–790, 2002.

[46] R.S. Thorne and R.G. Roberts. An Ordered analysis of heavy flavor production in deep inelastic scattering. *Phys.Rev.*, D57:6871–6898, 1998.

[47] R.S. Thorne. A Variable-flavor number scheme for NNLO. *Phys.Rev.*, D73:054019, 2006.

[48] S. Chekanov et al. An NLO QCD analysis of inclusive cross-section and jet-production data from the zeus experiment. *Eur.Phys.J.*, C42:1–16, 2005.

[49] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, Pavel M. Nadolsky, et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 0207:012, 2002.

[50] Serguei Chatrchyan et al. Measurement of the Inclusive Jet Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV. *Phys.Rev.Lett.*, 107:132001, 2011.

[51] Serguei Chatrchyan et al. Measurements of differential jet cross sections in proton-proton collisions at sqrt s =7 TeV with the CMS detector. 2012.

[52] J. Rojo. PDF Meeting at CMS. `https://indico.cern.ch/getFile.py/access?contribId=5&resId=0&materialId=slides&confId=194393`. Date: 09.07.2012.

[53] Serguei Chatrchyan et al. Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS. *JINST*, 6:P11002, 2011.

[54] Jet Energy Resolution and Corrections. Description of jet energy scale uncertainty sources, `"https://twiki.cern.ch/twiki/bin/viewauth/CMS/JECUncertaintySources"`.

[55] Andrea Banfi, Gavin P. Salam, and Giulia Zanderighi. Phenomenology of event shapes at hadron colliders. *JHEP*, 1006:038, 2010.

[56] V.M. Abazov et al. Determination of the strong coupling constant from the inclusive jet cross section in $p\bar{p}$ collisions at sqrt(s)=1.96 TeV. *Phys.Rev.*, D80:111107, 2009.

[57] F. Stober. *Determination of the Three-Jet mass cross section at sqrt(s)=7 TeV*. PhD thesis, University of Karlsruhe, 2012. IEKP-KA/2012-24.

[58] Measurement of the ratio of the inclusive 3-jet to 2-jet cross-sections in pp collisions at 7 TeV and first determination of the strong coupling at transverse momenta in the TeV range. 2011.

[59] M. Gluck, E. Reya, and A. Vogt. Dynamical parton distributions revisited. *Eur.Phys.J.*, C5:461–470, 1998.

[60] S. Bethke. Determination of the QCD coupling α_s. *J.Phys.*, G26:R27, 2000.

[61] T. Affolder et al. Measurement of the strong coupling constant from inclusive jet production at the Tevatron $\bar{p}p$ collider. *Phys.Rev.Lett.*, 88:042001, 2002.

Danksagung

Ich möchte mich sehr herzlich bei Prof. Günter Quast für die Möglichkeit, diese Studien durchzuführen sowie für die hervorragende Betreuung während meiner gesamten Zeit am EKP bedanken. Ebenso möchte ich mich bei Prof. Ulrich Husemann für die Übernahme des Korreferats bedanken.

Mein besonderer Dank gilt Dr. Klaus Rabbertz für die Betreuung und Korrektur dieser Diplomarbeit. Durch sein unglaubliches Wissen auf dem Feld der QCD hat er diese Arbeit erst möglich gemacht. Des Weiteren möchte ich mich bei Joram Berger, Fred Stober und Manuel Zeise für die vielen Diskussionen über Physik und das Korrekturlesen meiner Arbeit bedanken.

Ohne die Veröffentlichung der Software HERAFitter sowie die Unterstützung bei der Einarbeitung durch die Entwickler wären diese Studien nicht durchführbar gewesen. Insbesondere Ringaile Placakyte und Krzysztof Nowak bin ich sehr dankbar für ihre Hilfe und Anregungen.

Darüber hinaus gilt mein Dank meinen Kollegen Max Fischer, Raphael Friese, Corinna Günth, Dominik Haitz, Oliver Oberst und Andreas Oehler für die sehr angenehme Atmosphäre in der Arbeitsgruppe.

Meinen Eltern danke ich für ihre vielseitige Unterstützung und ihr Vertrauen während meines gesamten Studiums.

Hiermit versichere ich, die vorliegende Arbeit selbstständig verfasst und nur die angegebenen Hilfsmittel verwendet zu haben.

Georg Sieber

Karlsruhe, den 15. März 2013