



Unbiased Predictions for LHC Processes with HQ Mass Effects

Advisor: Prof. José Ignacio LATORRE
Co-advisor: Doc. Juan ROJO CHACÓN

Master Thesis of:
Francesco CERUTTI
NIUB 15415993

Academic Year 2009-2010

Contents

1	The NNPDF Approach	9
2	Heavy Quarks Implementation: the FONLL Method	14
3	Coefficient Functions in the FONLL Scheme	19
3.1	Mellin Space Results for HQ Neutral Current Massive Coefficient Functions	20
3.2	Mellin Space Results for HQ Charged Current Massive Coefficient Functions	23
4	Benchmarks and Impact on LHC Physics	26
5	Conclusions and Outlook	35
A	HQ CC Massive Coefficient Functions in x and N Mellin Space	38

Introduction

QCD and Factorization

The theory of strong interaction (QCD), as we know it today, is the result of a long theoretical and phenomenological evolution. After that a huge amount of new particles were observed in the 50s, a new theoretical description of some underlying and more fundamental structure was needed. In 1964 Gell-Mann and Zweig introduced the idea of new fundamental particles, the quarks, and in 1965 Han, Nambu, and Greenberg proposed an additional SU(3) gauge degree of freedom, due to a new quantum number: the color charge. These are the basis of quantum chromodynamics. This theory soon revealed one of its main properties: the quark confinement. It seems impossible to observe directly coloured particles because the strength of the interaction with other partons (as R.P.Feynman called quarks and gluons) increases as we try to extract a single quark from an hadron (or a meson). The coupling constant of strong interaction increases at low energies, making impossible to perform a perturbative approach as with QED.

Only in recent years, after the work on factorization due to D.Gross, D.Politzer, and F.Wilczek was published [1,2], has been possible to compute precise quantitative predictions for many high energy processes. They analyzed deep-inelastic lepton-hadron scattering in asymptotically free gauge theories of the strong interactions. Considering the problematic part of the squared amplitude, that is the product of two hadronic currents inside the hadronic tensor, it is possible to expand this product employing Wilson's operator product expansion. The dominant operators in the kinematic region considered are those of twist (dimension minus spin) two. This allows us to separate the non-perturbative contributions from the perturbative ones.

So we can write:

$$\langle p|J^\mu(x)J^\nu(0)|p\rangle \underset{|x|\rightarrow 0}{=} \sum_N C_N(x)\langle p|O_N^{\mu\nu}|p\rangle \quad (1)$$

and we can directly relate the moments of the structure functions with Wilson

coefficients and the matrix elements of the operator product expansion like this:

$$\int_0^1 dx x^{N-2} F_2(x, Q^2) = C_N \left(\frac{Q^2}{\mu^2}, \alpha_s(\mu^2) \right) A_N(\mu^2) \quad (2)$$

where

$$A_N = \langle p | O_N | p \rangle$$

and x , Q^2 , and μ^2 are respectively the proton's momentum fraction, the energy scale, and the factorization scale. Although in the original paper the result is found for deep-inelastic scattering processes, this method applies more generally to other hard-scattering scenarios.

Here we can read the quantities A_N and C_N as

$$A_N = \int_0^1 dx x^{N-1} q(x) \quad C_N = \int_0^1 dx x^{N-1} \tilde{C}(x) \quad (3)$$

and so, applying the convolution theorem, we can write a relation between the structure function in the x Mellin space and the functions q and \tilde{C} :

$$F_2(x, Q^2) = x \sum_{i=q, \bar{q}, g} \int_x^1 \frac{dy}{y} \tilde{C}_i(y, \alpha_s(Q^2)) q_i\left(\frac{x}{y}, Q^2\right) \quad (4)$$

This relation is of great importance because we can directly relate a physical observable (measured from the experiment) and our theoretical computation. In particular, the \tilde{C}_i functions represents the partonic or hard-scattering process and can be computed in perturbation theory while the q_i functions must be extracted from experimental datas, as they contain all the non-perturbative elements. They are the key to calculations involving hadrons and are usually called Parton Distribution Functions (PDF).

Parton Distribution Functions

As the idea of factorization was introduced, first attempts to find a set of parton distribution functions were produced. At a very early stage [3–8] PDFs were determined through a combination of general physical principles (sum rules for example), model assumptions, and the first crude experimental information coming from Bjorken scaling and its violation. The main aim of these first semi-quantitative determinations was to show the compatibility of the data with the partonic interpretation of hard-processes. In this way it was possible to perform tests of the theory of strong interactions [4–8].

In the 80s, as the accuracy of the data and the confidence in perturbative QCD improved, the gluon distribution was extracted from scaling violations [9], and first parton sets based on consistent global fits were performed [10, 11]. In that decade these sets were largely used for phenomenology and even if next-to-leading order tools were already available [12] the leading order was

considered accurate enough. The evolution to a next-to-leading order approach was claimed by a second generation of high-precision experiments (deep-inelastic scattering [13] and hadron collider experiments [14]) that led QCD towards precision physics. Next-to-leading order parton sets became standard analysis tools and were constantly updated during the ensuing decade [15]. A great improvement was due to HERA collider experiments [16]. The data collected at HERA increased substantially the kinematical region covered, improving the accuracy in parton sets (some known sets are CTEQ5 [17] and MRST2001 [18]) and in the determination of quantities which are sensitive to scaling violations.

Each group that produces parton sets performs fits assuming as functional form for the parton parametrization a standard form: $f(x) \sim x^\alpha(1-x)^\beta$. The differences between the parton sets of each group are given by the many technical details that follow the choice of the parametrization. For almost fifteen years parton sets were produced without errors. Even if the problem was rised and discussed since the beginning, only by the end of the 80s the sets of Ref. [19] included error estimates. The technique used was quite unsatisfactory as it consisted in comparing results obtained with many different parton sets. This method cannot perform a reliable estimation of systematic errors, as they would affect each parton set in a similar way. Many issues rised about the determination of reliable uncertainties for parton distributions as is the case of how to treat correlated uncertainties of underlying data.

An entirely different approach was suggested in the seminal papers Ref. [20,21]. Stressing the need for a systematic way to handle properly the problem, a combined solution based on Bayesian inference and Monte Carlo methods was proposed. The actual accuracy in data and the high precision needed in recent observable calculations have made compulsory the use of PDFs with reliable uncertainties. To this extent, choosing a fixed functional form for parton fits doesn't seem to generate sufficiently flexible fits in order to avoid systematic bias.

In the LHC era, to produce parton sets from global fits with the highest possible accuracy is a really important goal. To this aim is important to implement in PDFs determination all the tools that can improve the precision of the results computed with parton sets (as NNLO computations, massive treatment of heavy quarks, resummation). A great effort is now spent on the best way to combine the parton sets produced by different groups (CTEQ, MSTW, NNPDF...) and on error treatment. This is one of the aims of PDF4LHC working group, started in 2008 and actually ongoing to provide guidance on PDF to LHC experiments and phenomenology. Once the parton sets will reach the state-of-the-art on the bunch of old datasets will be possible to perform precision measurements at LHC and improve again PDFs accuracy including in global fits new data from LHC experiments.

In this thesis work I'll discuss the implementation within the NNPDF framework of heavy quark mass effects. In particular I'll focus on coefficient functions implementation for neutral and charged current observable in the FONLL scheme, their benchmarking, and their impact on PDFs and LHC physics.

Chapter 1

The NNPDF Approach

The NNPDF (Neural Network Parton Distribution Functions) approach is a brand new method to determine PDF sets without using a fixed functional form as parametrization. The main ingredient is the use of neural networks as basic interpolating functions combined with a Monte Carlo method. The complexity of the problem can be understood if we think that we are trying to find the functional form of parton distribution functions, that is an infinite-dimensional object, starting from a finite set of experimental points. Unless we make some extra assumptions, this is clearly an ill-posed problem. Usually this is exactly the strategy: to solve the problem assuming that the PDFs can be described with some functional form, parametrized by a finite-dimensional set of parameters that can be fitted from the data. In this way we solve the problem but at the prize of a loss in generality and flexibility. As we are choosing a particular functional form and excluding all the others we are easily introducing a bias in our fit.

With the neural network approach the problem is elegantly solved without the choice of a precise functional form. The first step is to generate a Monte Carlo sampling of data space. For each experimental point a sample of N_{rep} replicas of the point is generated around the original value with the probability distribution of the data and N_{rep} is large enough to reproduce the statistical properties of the data to the desired accuracy. For each replica of the data is performed a global fit so that if we produce 100 MC replicas we'll find 100 PDF replicas. Each PDF is parametrized by an individual neural network with 37 parameters so for each set we have $37 \times 7 = 259$ parameters. Fitting a complex shape requires a large number of parameters and so longer fitting. The large set of parameters guarantees the independence of the results on the parametrization. For each replica the parameters are randomly initialized and also strongly reducing their number the results are hardly affected. The efficiency of the neural network technology is due to its nonlinear behaviour. Even if the name could

sound as referred to some highly complicated object, a neural network is quite easy to build. We use a multilayer feed-forward network: each neuron receives input from neurons in the preceding layer and feeds output to neurons in the subsequent layer. Each neuron is defined by an activation function, a sigmoid:

$$g(x) = \frac{1}{1 + e^{-\beta x}} \quad (1.1)$$

and the activation is determined by weights and thresholds inside the argument of the sigmoid:

$$\xi_i = g\left(\sum_j \omega_{ij} \xi_j - \theta_i\right) \quad (1.2)$$

So, for a neural network with architecture 1-2-1 we can write its functional form as:

$$f(x) = \frac{1}{1 + e^{\theta_1^{(3)} - \frac{\omega_{11}^{(2)}}{1 + e^{\theta_1^{(2)} - x\omega_{11}^{(1)}}} - \frac{\omega_{12}^{(2)}}{1 + e^{\theta_2^{(2)} - x\omega_{21}^{(1)}}}}} \quad (1.3)$$

Physical observables are computed from parton distribution functions: the computed PDF is evolved from the initial scale to the scale at which data are available by using standard QCD evolution equations and physical observables are computed by convoluting the evolved parton distributions with hard partonic cross sections. The optimal fit is not determined when the χ^2 reaches its minimum value but is determined by a cross-validation method. Data are splitted in a training set and a validation set. Replicas are fitted to the training subset (each replica on a different random subset) and the χ^2 is computed both with the validation and with the training subset. When the χ^2 on the validation data stops improving, the parameters of the neural network are stored. In this way overlearning is avoided. In fact a too long training time leads to a fit also of data's statistical fluctuations and not only of the underlying distribution. PDFs behave as expected upon the addition of new data (e.g. uncertainties expand when data are removed and shrink when they are added unless the new data is incompatible with the old) [22–27] and results are even stable upon the addition of new independent PDFs [22–25].

Over the last several years the NNPDF Collaboration has developed this novel approach to the determination of parton distribution functions. The method was developed, refined, and applied to problems of increasing complexity: the parametrization of a single structure function [28], of several structure functions [29] and the determination of the nonsinglet parton distribution [30]. Eventually, in Ref. [22] a first complete set of parton distributions was constructed, using essentially all the then-available deep-inelastic scattering (DIS) data. This parton set, NNPDF1.0, includes five independent parton distributions (the two lightest flavours and antiflavours and the gluon). It was then

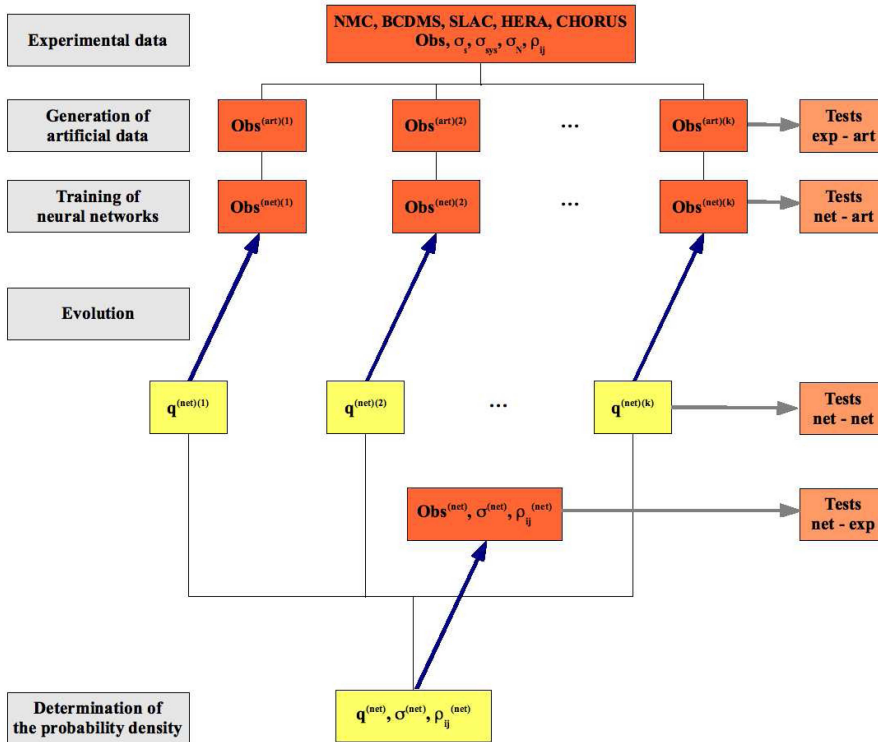


Figure 1.1: Schematic representation of the NNPDF approach.

extended in Refs. [25,26] to also include an independent parametrization of the strange and antistrange quarks, with heavier flavours determined dynamically (NNPDF1.2 parton set). All NNPDF parton sets are available through the LHAPDF interface [23,24]. In these works, as well as in studies for the HERA-LHC workshop [27] it was shown that PDFs determined using the NNPDF methodology enjoy the several features already explained above. With PDF uncertainties under control, detailed precision physics studies become possible, such as for instance the determination of CKM matrix elements [26].

It has been known for a long time (see Ref. [22] for references to the earlier literature) that DIS data are insufficient to determine accurately many aspects of PDFs, such as the flavour decomposition of the quark and antiquark sea or the gluon distribution, especially at large x : indeed, the current state-of-the-art PDF determinations, such as the CTEQ6.6 [31] and MSTW2008 [32] are based

on global fits, in which hadronic data are included along with DIS data. In the last work (NNPDF2.0) a PDF determination using NNPDF methodology based on a global fit is presented. The data used for fitting include, on top of all the data used in Ref. [26] (DIS data and “dimuon” charm neutrino production data) also hadronic data, specifically Drell-Yan (DY), W and Z production and Tevatron jets. We also replace the separated ZEUS and H1 datasets with the recently published HERA-I combined dataset [33]. The dataset used in this parton determination is thus comparable in variety and size (and is in fact slightly larger) to that used by the CTEQ and MSTW groups.

The PDF determination presented with the NNPDF2.0 set is based on a consistent use of NLO QCD. This is novel in the context of a global parton determination: indeed, in other parton fits such as Refs. [31, 32] only DIS data are treated using fully NLO QCD, while several sets of hadronic data are treated using LO theory improved through K-factors.

The main bottleneck in the use of NLO theory for hadronic processes is the speed in the computation of hadron-level observables, which requires a convolution of the PDF of both incoming hadrons with parton-level cross sections. The use of Mellin-space techniques (as e.g. in Ref. [34]) solves this problem, but at the cost of limiting the flexibility of the acceptable PDF parametrization: specifically, the very flexible neural network method of Refs. [22–26] parametrizes PDFs in x space. Efficient fast methods to overcome this hurdle have been suggested (see [35], in [36]), based on the idea of precomputing and storing the convolution with a set of basis functions over which any PDF can be expanded.

To extract the NNPDF2.0 parton set, we use similar ideas to fully exploit the powerful parton evolution method introduced in Ref. [30], based on the convolution of PDFs with a pre-computed kernel, determined using Mellin-space techniques. This gives us a new approach, which we call the FastKernel method, which we use both for parton evolution, and for the computation of DIS and DY physical observables. The FastKernel method leads to a considerable increase in speed in comparison to Refs. [22–26] for DIS data, and it makes possible for the first time to use exact NLO theory for DY in a global parton fit. Thanks to FastKernel, we are able to produce a first fully NLO global parton set using NNPDF methodology. This parton determination enjoys the same desirable features of the previous NNPDF1.0 and NNPDF1.2 PDF sets, with which in particular it is fully compatible, though uncertainties are now significantly smaller, and in fact sometimes also rather smaller than those of previous global fits.

Thanks to the use of a Monte Carlo methodology, it is possible to perform a detailed comparison of NNPDF2.0 with those of previous NNPDF fits, and in particular to assess the impact of the various new aspects of this parton determination, both coming from improved methodology and the use of more precise data and a wider dataset. The most striking feature of the NNPDF2.0 parton

determination is perhaps the fact that it is free of tension between different data sets and NLO QCD: in fact, whereas the addition of new data leads to sizable error reduction, essentially no data incompatibility is found, and the distribution of the fit results appears to be fully consistent with statistical expectations.

The NNPDF collaboration is actually working on a new parton set: NNPDF2.1. Here heavy quarks effects are implemented into the FastKernel framework and the PDFs are determined from a global set of hard scattering data which include all relevant experimental data on the charm structure functions in deep-inelastic scattering from HERA experiments.

Chapter 2

Heavy Quarks Implementation: the FONLL Method

Only in very recent years heavy quark effects started being considered. The main reason that called for a proper treatment of these effects was the discovery [37] that mass-suppressed terms in global fits can affect predictions for the total W and Z production at the LHC by almost 10%. Before then, in many global fit analysis concerning PDFs, heavy quarks were supposed to decouple from the structure functions for $Q^2 < m_h$ and were considered massless for $Q^2 > m_h$.

A technique for the inclusion of heavy mass-suppressed contributions to structure functions was developed long ago [38, 39], based upon a renormalization scheme with explicit heavy quark decoupling [40] (called also CWZ scheme). Several variants of this method (usually called ACOT) were subsequently proposed, such as S-ACOT [41] and ACOT- χ [42, 43]. However, the ACOT method was first used for an actual global parton fit only recently. An alternative method (sometimes called TR) has also been advocated [44–46], and used for parton fits. Recently, however, the methods used by the CTEQ and MSTW groups for their current parton fits, based respectively on the ACOT and TR procedures, have adopted at least in part a common framework: they have been compared recently in Refs. [47, 48], thereby elucidating differences and common aspects.

The FONLL method was first introduced in Ref. [49] in the context of hadroproduction of heavy quarks. The name of this method comes from the original work, where a fixed order calculation (second order) was combined with a next-to-leading log one. But the method is general and can be used consistently to combine a fixed order with a resummed calculation to any or-

der of either. The FONLL method only relies on standard QCD factorization and calculations with massive quarks in the CWZ decoupling scheme and with massless quarks in the $\overline{\text{MS}}$ scheme. As ACOT and TR procedures, FONLL gives a new way of matching the decoupling scheme ($Q^2 \leq m_h^2$) with the $\overline{\text{MS}}$ scheme. In fact the basic problem in the treatment of heavy quarks in a QCD process stems from the fact that QCD calculations are usually performed in a decoupling scheme, rather than in the $\overline{\text{MS}}$ scheme.

Indeed, in the standard $\overline{\text{MS}}$ scheme, heavy quark contributions are present at all scales: for instance, the β function depends on n_f with $n_f = 6$ at all scales. In a decoupling scheme, instead, in the computation of a process characterized by the hard scale Q^2 all quarks with mass $m_q^2 > Q^2$ (heavy, henceforth) are not treated in the $\overline{\text{MS}}$ scheme. Rather, heavy flavour graphs are subtracted at zero momentum. The important consequence of this definition is that heavy quarks decouple for scales much lower than the heavy quark mass. This implies that the GLAP evolution equations, and the running of α_s are identical to those which one would get in the $\overline{\text{MS}}$ scheme, but with n_f equal to the number of light flavours, $n_f = n_l$. If for $Q^2 < m_h^2$ one neglects all terms suppressed by powers of Q^2/m_h^2 , one obtains a so-called zero-mass scheme, where all quarks are treated as massless, but heavy quarks are absent at $Q^2 \leq m_h^2$. This scheme can be combined with the usual $\overline{\text{MS}}$ scheme for $n_l + 1$ flavours, including the heavy quark, when $Q^2 \geq m_h^2$; if terms suppressed by powers of m_h^2/Q^2 are neglected throughout this yields the so called zero-mass variable-flavour number scheme (ZMVFN from now on).

The ZMVFN scheme is not accurate near the threshold region, where $\frac{m_h^2}{Q^2} \sim 1$. The problem in this region is easily remedied by simply using the decoupling scheme with $n_f = n_l$, but retaining explicitly the full dependence on the heavy quark mass in the computation of hard cross sections. This way of computing however loses accuracy in comparison to the previous ZMVFN when $Q^2 \gg m_h^2$ so that $L \equiv \ln Q^2/m_h^2 \gg 1$, because, in the decoupling scheme with $n_f = n_l$, these large logs are only included to fixed order in α_s while in the ZMVFN scheme they are resummed to all orders.

While in both the ACOT and TR schemes the matching conditions ensue from the requirement that computations of the same observable within different renormalization schemes give the same answer within the respective accuracy, the FONLL scheme is based on a simpler idea: the decoupling and ZMVFN schemes are combined and double counting terms between the two computations are subtracted order by order in an expansion in powers of α_s and L .

Let's consider as an example a generic structure function F . This observable will have two different expressions, one in the decoupling scheme and another one in the ZMVFN scheme. Both these expressions factorize in terms of coefficient functions and PDFs, but in the first case, that we can also call massive,

the coefficient functions are computed fully retaining the mass dependence and $\alpha_s^{(n_l)}$ and the PDFs $f_i^{(n_l)}(y, Q^2)$ obey standard $\overline{\text{MS}}$ evolution equations with n_l flavours. In the second case instead the expressions are fully massless. The strong coupling and the PDFs satisfy standard GLAP equations with $n_f = n_l + 1$ and the PDFs include the heavy flavour as a light parton.

In order to carry out the FONLL procedure, we need to express the decoupling scheme structure function in terms of $\alpha_s^{(n_l+1)}$ and the PDFs $f_i^{(n_l+1)}$ (that, as the $n_l + 1$ suggests, are in the massless scheme) for $i \neq h, \bar{h}$. The coupling constant and PDFs are related in the two schemes by equations of the form

$$\alpha_s^{(n_l+1)}(Q^2) = \alpha_s^{(n_l)}(Q^2) + \sum_{i=2}^{\infty} c_i(L) \times \left(\alpha_s^{(n_l)}(m^2) \right)^i \quad (2.1)$$

$$f_i^{(n_l+1)}(x, Q^2) = \int_x^1 \frac{dy}{y} \sum_{j=q, \bar{q}, g} K_{ij} \left(\frac{x}{y}, L, \alpha_s^{(n_l)}(Q^2) \right) f_j^{(n_l)}(y, Q^2) \quad (2.2)$$

where

$$L \equiv \log Q^2/m^2. \quad (2.3)$$

The coefficients $c_i(L)$ are polynomials in L , and the functions K_{ij} can be expressed as an expansion in powers of α_s , with coefficients that are polynomials in L . From these matching conditions we can, inverting them, recover an expression for the massive scheme structure function in terms of $\alpha_s^{(n_l+1)}$ and the PDFs $f_i^{(n_l+1)}$:

$$F^{(n_l)}(x, Q^2) = x \int_x^1 \frac{dy}{y} \sum_{i=q, \bar{q}, g} B_i \left(\frac{x}{y}, \frac{Q^2}{m^2}, \alpha_s^{(n_l+1)}(Q^2) \right) f_i^{(n_l+1)}(y, Q^2) \quad (2.4)$$

where the coefficient functions B_i are such that substituting the matching relations (2.1-2.2) in $F^{(n_l)}$ we recover the original expression in the decoupling scheme.

In order to match the two expressions for F (in the massive and massless schemes) we work out their perturbative expansion. Thanks to GLAP equations and to Eq.(2.2) we can express, in the absence of intrinsic heavy quark contributions, the heavy quark PDFs present in the massless-scheme expression of F in terms of the light-quark PDFs. Thus, the massless-scheme expression of F may be written entirely in terms of light-quark PDFs:

$$F^{(n_l+1)}(x, Q^2) = x \int_x^1 \frac{dy}{y} \sum_{i=q, \bar{q}, g} A_i^{(n_l+1)} \left(\frac{x}{y}, L, \alpha_s^{(n_l+1)}(Q^2) \right) f_i^{(n_l+1)}(y, Q^2) \quad (2.5)$$

where note that the sum is performed only on light flavours and the $A_i^{(n_l+1)}$

coefficient functions are given by a perturbative expansion of the form

$$A_i^{(n_l+1)}\left(z, L, \alpha_s^{(n_l+1)}(Q^2)\right) = \sum_{p=0}^N \left(\alpha_s^{(n_l+1)}(Q^2)\right)^p \sum_{k=0}^{\infty} A_i^{p,k}(z) \left(\alpha_s^{(n_l+1)}(Q^2)L\right)^k \quad (2.6)$$

where at leading order $N = 0$, and at N^kLO $N = k$.

Also the massive expression of F is written in terms of light-quark PDFs, and we can write a finite-order perturbative expansion for the B_i massive coefficient functions:

$$B_i\left(z, \frac{Q^2}{m^2}, \alpha_s^{(n_l+1)}(Q^2)\right) = \sum_{p=0}^P \left(\frac{\alpha_s^{(n_l+1)}(Q^2)}{2\pi}\right)^p B_i^p\left(z, \frac{Q^2}{m^2}\right) \quad (2.7)$$

where P is the order of the expansion needed to reach the desired accuracy.

Now we have all the ingredients to expose in a simply way the main idea of the FONLL method: in the massless scheme expression Eq.(2.5) one replaces all contributions to the expansion (2.6) of the coefficient functions $A_i^{(n_l+1)}\left(z, L, \alpha_s^{(n_l+1)}(Q^2)\right)$ which appear in

$$B_i^{(0),p}\left(x, \frac{Q^2}{m^2}\right) \equiv \sum_{k=0}^p A_i^{p-k,k}(x)L^k \quad (2.8)$$

where

$$\lim_{m \rightarrow 0} \left[B_i^p\left(x, \frac{Q^2}{m^2}\right) - B_i^{(0),p}\left(x, \frac{Q^2}{m^2}\right) \right] = 0 \quad (2.9)$$

with their fully massive expression $B_i^p\left(x, \frac{Q^2}{m^2}\right)$ from the massive scheme expression of F . In this way all mass suppressed effects that are not present in the massless expression but are known from the massive one are included. At a structure function level, this means that we first of all need to define a massless limit of the massive-scheme expression Eq.(2.4), namely

$$F^{(n_l,0)}(x, Q^2) = x \int_x^1 \frac{dy}{y} \sum_{i=q,\bar{q},g} B_i^{(0)}\left(\frac{x}{y}, \frac{Q^2}{m^2}, \alpha_s^{(n_l+1)}(Q^2)\right) f_i^{(n_l+1)}(y, Q^2) \quad (2.10)$$

where

$$B_i^{(0)}\left(z, \frac{Q^2}{m^2}, \alpha_s^{(n_l+1)}(Q^2)\right) = \sum_{p=0}^P \left(\frac{\alpha_s^{(n_l+1)}(Q^2)}{2\pi}\right)^p B_i^{(0),p}\left(z, \frac{Q^2}{m^2}\right) \quad (2.11)$$

and then we have to put things together in the right way to obtain the FONLL expression of our observable:

$$F^{\text{FONLL}}(x, Q^2) = F^{(d)}(x, Q^2) + F^{(n_l)}(x, Q^2) \quad (2.12)$$

with

$$F^{(d)} \equiv \left[F^{(n_l+1)}(x, Q^2) - F^{(n_l,0)}(x, Q^2) \right] \quad (2.13)$$

where the expressions $F^{(n_i)}$, $F^{(n_i+1)}$, and $F^{(n_i,0)}$ are given respectively by Eqs.(2.4), (2.5), and (2.10).

Thanks to Eq.(2.9) we can see that when $Q^2 \gg m^2$ the massive-scheme term $F^{(n_i)}(x, Q^2)$ tends to $F^{(n_i,0)}(x, Q^2)$ and so the FONLL expression (2.12) reduces to the massless-scheme expression $F^{(n_i+1)}(x, Q^2)$. When instead $Q^2 \approx m^2$ we have that $F^{(n_i,0)}(x, Q^2)$ kills in $F^{(n_i+1)}(x, Q^2)$ what is already contained inside $F^{(n_i)}(x, Q^2)$ and, up to mass suppressed terms, we recover the $F^{(n_i)}(x, Q^2)$ expression. Here a mismatch in accuracy can occur. In facts to avoid this to happen we have to compute $F^{(n_i+1)}$ with an accuracy at least as high as that of $F^{(n_i)}$ (and, as a consequence, as that of the massless limit $F^{(n_i,0)}$). If this doesn't happens we have that with $F^{(n_i,0)}$ we are subtracting to $F^{(n_i+1)}$ terms that are not present in its expansion. However it's possible to generalize Eq.(2.12) in such a way that it can still be used even when $F^{(n_i)}$ is known with higher precision than $F^{(n_i+1)}$.

For this purpose it is sufficient to retain in $F^{(n_i,0)}$ only those terms that are also present in $F^{(n_i+1)}$. Following this strategy we indeed loose accuracy in the $Q^2 \gg m^2$ limit because $F^{(n_i,0)}$ and $F^{(n_i)}$ no more cancel in this limit. In this case we only can say that the FONLL method reduces to the massless scheme expression $F^{(n_i+1)}$ up to mass suppressed terms and terms of higher order in $\mathcal{O}(\alpha_s^2)$.

As last issue of this chapter, it should be observed that the $F^{(d)}$ term, even if subleading, could in practice be non-negligible near the threshold m_h^2 . For this reason we choose to multiply it by a damping factor of the form

$$\mathcal{D}(Q^2) = \theta(Q^2 - m_h^2) \left(1 - \frac{m_h^2}{Q^2}\right)^2. \quad (2.14)$$

We perform this choice because near threshold the difference factor could become totally unreliable, and it is possible to suppress $F^{(d)}$ without modifying the accuracy of our computation just because it is a subleading term.

Chapter 3

Coefficient Functions in the FONLL Scheme

For our purpose it's now better to write the FONLL observables separating off explicitly the light and heavy contribution like this:

$$F_l^{\text{FONLL}}(x, Q^2) = F_l^{(n_l)}(x, Q^2) + F_l^{(d)}(x, Q^2) \quad (3.1)$$

$$F_h^{\text{FONLL}}(x, Q^2) = F_h^{(n_h)}(x, Q^2) + F_h^{(d)}(x, Q^2) \quad (3.2)$$

which is just a consequence of writing the generic structure function $F(x, Q^2)$ and the generic coefficient function $C_i(x, \alpha_s(Q^2))$ in the same splitted way:

$$F(x, Q^2) = F_l(x, Q^2) + F_h(x, Q^2) \quad (3.3)$$

$$C_i(x, \alpha_s(Q^2)) = C_{i,l}(x, \alpha_s(Q^2)) + C_{i,h}(x, \alpha_s(Q^2)) \quad (3.4)$$

where we define F_h and $C_{i,h}$ as the contributions to F and C_i respectively which are obtained when only the electric charge of the heavy quark is nonzero. Note that therefore, up to $\mathcal{O}(\alpha_s^2)$, in the coefficient functions the label l, h denotes the quark to which the virtual photon couples, whereas the label i denotes the parton that enters the hard scattering process.

While the matching condition for parton distributions at the scale $Q^2 = m^2$, is trivially

$$f_i^{(n_i+1)}(x, m^2) = f_i^{(n_i)}(x, m^2) + \mathcal{O}(\alpha_s^2) \quad (3.5)$$

and is sufficient in order to fix the relation between PDFs in the massive and massless schemes, in order to implement the FONLL approximation we need to express the massive-scheme PDFs in terms of the massless-scheme ones at a generic scale Q^2 , as required in Eq. 2.4. The relation at any scale, for a calculation at order α_s , can be obtained from the matching condition Eq. 3.5

by solving the evolution equation for both $f_i^{(n_i+1)}$ and $f_i^{(n_i)}$ in the respective schemes.

In this thesis work the coefficient function's computation is considered only to order α_s , so the matching condition for PDFs at any scale Q^2 starts at order α_s and the correction is non-zero only for the gluon PDF:

$$f_g^{(n_i+1)}(x, Q^2) = f_g^{(n_i)}(x, Q^2) - \frac{\alpha_s(Q^2)}{2\pi} \frac{2T_R}{3} L f_g^{(n_i+1)}(x, Q^2) + \mathcal{O}(\alpha_s^2) \quad (3.6)$$

As a consequence we have that the FONLL expression of the light part of the structure function, F_l^{FONLL} , it reduces to the massless scheme part, $F_l^{(n_i+1)}$. The heavy component instead, given by equation 3.2, it reduces to the sum of

$$\begin{aligned} F_h^{(n_i)}(x, Q^2) &= x \int_x^1 \frac{dy}{y} C_{g,h}^{(n_i)}\left(\frac{x}{y}, \frac{Q^2}{m^2}, \alpha_s(Q^2)\right) f_g^{(n_i+1)}(y, Q^2) \quad (3.7) \\ F_h^{(d)}(x, Q^2) &= x \int_x^1 \frac{dy}{y} \left[C_{q,h}^{(n_i+1)}\left(\frac{x}{y}, \alpha_s(Q^2)\right) \left[f_h^{(n_i+1)}(y, Q^2) + f_{\bar{h}}^{(n_i+1)}(y, Q^2) \right] \right. \\ &\quad \left. + \left(C_{g,h}^{(n_i+1)}\left(\frac{x}{y}, \alpha_s(Q^2)\right) - B_{g,h}^{(0)}\left(\frac{x}{y}, \frac{Q^2}{m^2}, \alpha_s(Q^2)\right) \right) f_g^{(n_i+1)}(y, Q^2) \right] \quad (3.8) \end{aligned}$$

Noting that $C_{q,h}(x, \alpha_s(Q^2)) = \delta(1-x) + \mathcal{O}(\alpha_s)$ and that

$$f_h(y, Q^2) = f_{\bar{h}}(y, Q^2) = \frac{\alpha_s^2(Q^2)}{2\pi} L \int \frac{dz}{z} T_f(z^2(1-z)^2) g\left(\frac{x}{y}, Q^2\right) + \mathcal{O}(\alpha_s^2) \quad (3.9)$$

it's easy to see that, in the region where L is not large, the heavy FONLL expression coincides with the massive-scheme part (at order \mathcal{O}).

In the following sections are presented the heavy coefficient functions and their Mellin transforms for heavy quark production to order $\mathcal{O}(\alpha_s)$.

3.1 Mellin Space Results for HQ Neutral Current Massive Coefficient Functions

The gluon coefficient function in the massive (n_l) scheme is given in the FONLL notation by

$$C_{g,h}^{(n_l)}\left(z, \frac{Q^2}{m^2}, \alpha_s(Q^2)\right) = \frac{\alpha_s(Q^2)}{2\pi} C_{g,h}^{(n_l),1}\left(z, \frac{Q^2}{m^2}\right) + \mathcal{O}(\alpha_s^2) \quad (3.10)$$

where the leading $\mathcal{O}(\alpha_s)$ term is given by

$$\begin{aligned} C_{g,h}^{(n_l),1}\left(z, \frac{Q^2}{m^2}\right) &= \theta(W^2 - 4m^2) \times T_R [(z^2 + (1-z)^2 + 4\epsilon z(1-3z) - 8\epsilon^2 z^2) \log \frac{1+v}{1-v} \\ &\quad + (8z(1-z) - 1 - 4\epsilon z(1-z))v], \quad (3.11) \end{aligned}$$

where we have used the definitions

$$\epsilon \equiv m^2/Q^2, \quad v \equiv \sqrt{1 - 4m^2/W^2}, \quad W^2 \equiv Q^2 \frac{1-z}{z} \quad (3.12)$$

The Mellin transform of Eq. 3.11 is defined as

$$C_g^{(n_l),1} \left(N, \frac{Q^2}{m^2} \right) = \int_0^{(1+4m^2/Q^2)^{-1}} dz z^{N-1} C_g^{(n_l),1} \left(z, \frac{Q^2}{m^2} \right). \quad (3.13)$$

The integral Eq. 3.13 can be written as follows

$$C_g^{(n_l),1} (N, \epsilon) = T_R a^N \int_0^1 dt t^{N-1} \left\{ [1 + 2a(2\epsilon - 1)t + 2a^2(1 - 6\epsilon - 4\epsilon^2)t^2] \ln \frac{1+v}{1-v} - [1 + 4a(\epsilon - 2)t - 4a^2(\epsilon - 2)t^2] v \right\} \quad (3.14)$$

$$= T_R a^N \int_0^1 dt t^{N-1} \left\{ [1 + (1 - 3a)t - \frac{1}{2}(1 + 4a - 9a^2)t^2] \ln \frac{1+v}{1-v} - [1 + (1 - 9a)t - a(1 - 9a)t^2] v \right\} \quad (3.15)$$

where in the second line we used $4\epsilon a = 1 - a$, that is

$$a(\epsilon) = \frac{1}{1 + 4\epsilon}, \quad (3.16)$$

to simplify the coefficients. The integrals we need are thus

$$J_1(N) \equiv \int_0^1 dt t^{N-1} \ln \frac{1+v}{1-v} \quad J_2(N) \equiv \int_0^1 dt t^{N-1} v, \quad (3.17)$$

since extra powers of t can be accommodated by a shift in N by an integer. Here as usual $v = (1-t)^{1/2}/(1-at)^{1/2}$.

Now the key is that these two integrals are almost certainly related by an integration by parts, and since the second is doable, the first must be also. To show this, we need

$$\begin{aligned} \frac{d}{dt} \ln \frac{1+v}{1-v} &= \frac{dv}{dt} \frac{d}{dv} \ln \frac{1+v}{1-v} \\ &= \left(-\frac{1}{2} \frac{1-a}{(1-t)^{1/2}(1-at)^{3/2}} \right) \left(\frac{2(1-at)}{(1-a)t} \right) \\ &= -\frac{1}{t} \frac{1}{(1-t)^{1/2}(1-at)^{3/2}} \end{aligned} \quad (3.18)$$

Thus

$$J_1(N) = -\frac{1}{N} \int_0^1 dt t^N \frac{d}{dt} \ln \frac{1+v}{1-v} = \frac{1}{N} I(N), \quad (3.19)$$

where

$$I(N) \equiv \int_0^1 dt t^{N-1} (1-t)^{-1/2} (1-at)^{-1/2}. \quad (3.20)$$

Note that the boundary term in the integration by parts vanishes for all $Re N > 0$, and thus its analytic continuation vanishes for all N , so can be safely ignored.

Trivially $J_2(N) = I(N) - I(N+1)$. The integral $I(N)$ may be evaluated in the usual way in terms of a hypergeometric:

$$I(N) = \frac{\Gamma(N)\Gamma(\frac{1}{2})}{\Gamma(N+\frac{1}{2})} F\left(\frac{1}{2}, N, N+\frac{1}{2}; a\right). \quad (3.21)$$

Note that when $a = 0$, this reduces to $B(N, \frac{1}{2})$, as it should.

We thus have

$$\begin{aligned} C_g^{(n_l, 1)}(N, \epsilon) &= T_R a^N \left\{ \left[\frac{1}{N} I(N) + \frac{1-3a}{N+1} I(N+1) - \frac{1}{2} \frac{1+6a-9a^2}{N+2} I(N+2) \right] \right. \\ &\quad - [I(N) - I(N+1) + (1-9a)(I(N+1) - I(N+2)) \\ &\quad \left. - a(1-9a)(I(N+2) - I(N+3))] \right\} \\ &= T_R a^N \left\{ \left(\frac{1}{N} - 1 \right) I(N) + \left(\frac{1-3a}{N+1} + 9a \right) I(N+1) \right. \\ &\quad \left. - \left(\frac{1}{2} \frac{1+4a-9a^2}{N+2} - (1+a)(1-9a) \right) I(N+2) - a(1-9a) I(N+3) \right\}. \end{aligned} \quad (3.22)$$

The massless limit of the massive CF Eq. 3.11 is given by

$$C_{g,h}^{(n_l, 0), 1} \left(z, \frac{Q^2}{m^2} \right) = T_R \left[(z^2 + (1-z)^2) \log \frac{Q^2(1-z)}{m^2 z} + (8z(1-z) - 1) \right], \quad (3.23)$$

whose Mellin transform is

$$\begin{aligned} C_g^{(n_l, 0), 1} \left(N, \frac{Q^2}{m^2} \right) &\equiv \int_0^1 dz z^{N-1} C_{g,h}^{(n_l, 0), 1} \left(z, \frac{Q^2}{m^2} \right) \\ &= T_R \frac{(2+N+4N^2-N^3) + N(2+N+N^2) \left(\log \left(\frac{Q^2}{m^2} \right) - H_N \right)}{N^2(1+N)(2+N)} \end{aligned} \quad (3.24)$$

A cross-check of the Mellin transform of the massive CF Eq. 3.22 is to show that its massless limit coincides with the Mellin transform of the x-space $(n_l, 0)$ coefficient function, Eq. 3.24. To this purposes, we need to expand Eq. 3.22 near $a = 1$. Near $a = 1$, ie $\epsilon = 0$, we need to use the expansion (AS 15.3.10)

$$\begin{aligned} F\left(\frac{1}{2}, N, N+\frac{1}{2}; a\right) &= \frac{\Gamma(N+\frac{1}{2})}{\Gamma(\frac{1}{2})^2 \Gamma(N)^2} \sum_{n=0}^{\infty} \frac{\Gamma(n+\frac{1}{2}) \Gamma(N+n)}{(n!)^2} \\ &\quad [2\psi(n+1) - \psi(n+\frac{1}{2}) - \psi(N+n) - \ln(1-a)] (1-a)^n. \end{aligned} \quad (3.25)$$

so

$$\begin{aligned} I(N) &= \frac{1}{\Gamma(\frac{1}{2}) \Gamma(N)} \sum_{n=0}^{\infty} \frac{\Gamma(n+\frac{1}{2}) \Gamma(N+n)}{(n!)^2} \\ &\quad [2\psi(n+1) - \psi(n+\frac{1}{2}) - \psi(N+n) - \ln(1-a)] (1-a)^n. \end{aligned} \quad (3.26)$$

The $n = 0$ term then gives the $\ln \epsilon$ collinear divergence, which is subtracted by the massless coefficient function: as $\epsilon \rightarrow 0$

$$I(N) \sim \ln(1/4\epsilon) - 2\gamma_E - \psi\left(\frac{1}{2}\right) - \psi(N) + \mathcal{O}(\epsilon). \quad (3.27)$$

Substituting in Eq. 3.22 we obtain

$$C_g^{(n_i),1}(N, \epsilon) = T_R \frac{[-N^3 + 3N^2 - N(2 + N + N^2)(\ln \epsilon + \gamma_E + \psi(N))]}{N^2(1 + N)(2 + N)} + \mathcal{O}(\epsilon), \quad (3.28)$$

as expected: the coefficient of the singularity is precisely γ_{qg} . Therefore we have checked that the massless limit is properly reproduced,

$$C_g^{(n_i),1}(N, \epsilon) + \mathcal{O}(\epsilon) = C_g^{(n_i,0),1}(N, \epsilon) \quad (3.29)$$

with the massive asymptotic coefficient function given in Eq. 3.24.

3.2 Mellin Space Results for HQ Charged Current Massive Coefficient Functions

In this section we compute the analytical Mellin transforms of the massive heavy quark coefficient functions in charged current (CC) deep-inelastic scattering. The $\mathcal{O}(\alpha_s)$ expressions were computed in Ref. [50].

The charm structure functions in charged current scattering in the massive scheme read

$$\begin{aligned} \mathcal{F}_{i,h}^{(n_i)}(x, Q^2) &= s'(\xi, \mu^2) + \frac{\alpha_s(\mu^2)}{2\pi} \left\{ \int_{\xi}^1 \frac{d\xi'}{\xi'} \left[C_{i,h}^{(n_i)}(\xi', \mu^2, \lambda) s'\left(\frac{\xi}{\xi'}, \mu^2\right) + \right. \right. \\ &\quad \left. \left. + C_{i,g}^{(n_i)}(\xi', \mu^2, \lambda) g\left(\frac{\xi}{\xi'}, \mu^2\right) \right] \right\} = \\ &= s'(x, \mu^2) + \frac{\alpha_s(\mu^2)}{2\pi} \left\{ \int_x^1 \frac{dx'}{x'} \theta(W^2 - m_c^2) \left[C_{i,h}^{(n_i)}(x', \mu^2, \lambda) s'\left(\frac{x}{x'}, \mu^2\right) + \right. \right. \\ &\quad \left. \left. + C_{i,g}^{(n_i)}(x', \mu^2, \lambda) g\left(\frac{x}{x'}, \mu^2\right) \right] \right\} \quad (3.30) \end{aligned}$$

with

$$\begin{aligned} s' &\equiv |V_{cs}|^2 s + |V_{cd}|^2 \frac{d+u}{2} \\ \xi &= x \left(1 + \frac{m_c^2}{Q^2}\right) \end{aligned}$$

and $\mathcal{F}_1^c \equiv F_{1,h}^{(n_i)}$, $\mathcal{F}_2^c \equiv \frac{F_2^c}{2\xi}$, $\mathcal{F}_3^c \equiv \frac{F_3^c}{2}$.

For a generic structure function,

$$F(x, Q^2, \lambda) = \int_{x/\lambda}^1 \frac{dy}{y} C^{(n_l)}(y, \lambda, \alpha_s) q\left(\frac{x}{\lambda y}, Q^2\right) \quad (3.31)$$

it can be shown that its Mellin transform

$$F(N, Q^2, \lambda) = C^{(n_l)}(N, \lambda, \alpha_s) q(N, Q^2) \quad (3.32)$$

is given in terms of Mellin moments of the x -space coefficient functions in Eq.3.30

$$C^{(n_l)}(N, \lambda, \alpha_s) = \int_0^\lambda dz z^{N-1} C^{(n_l)}(z/\lambda, \lambda, \alpha_s) = \lambda^N \int_0^1 dz z^{N-1} C^{(n_l)}(z, \lambda, \alpha_s). \quad (3.33)$$

The x -space explicit expressions at $\mathcal{O}(\alpha_s)$ of the CC massive coefficient functions given in Ref. [50] are reported below, both for quarks and gluons.

$$\begin{aligned} C_{1,q}^{(n_l)} &= \frac{4}{3} \left\{ - \left(4 + \frac{1}{2\lambda} + \frac{\pi^2}{3} + \frac{1+3\lambda}{2\lambda} K_A \right) \delta(1-z) + \right. \\ &- \frac{(1+z^2) \ln z}{1-z} + (1+z^2) \left[\frac{2 \ln(1-z) - \ln(1-\lambda z)}{1-z} \right]_+ + (1-4z+z^2) \frac{1}{(1-z)_+} + \\ &\left. + (z-z^2) \frac{1}{(1-\lambda z)_+} + \frac{1}{2} \left[\frac{1-z}{(1-\lambda z)^2} \right]_+ \right\} \quad (3.34) \end{aligned}$$

$$\begin{aligned} C_{2,q}^{(n_l)} &= \frac{4}{3} \left\{ - \left(4 + \frac{1}{2\lambda} + \frac{\pi^2}{3} + \frac{1+3\lambda}{2\lambda} K_A \right) \delta(1-z) + \right. \\ &- \frac{(1+z^2) \ln z}{1-z} + (1+z^2) \left[\frac{2 \ln(1-z) - \ln(1-\lambda z)}{1-z} \right]_+ + K_A \delta(1-z) + \\ &\left. + (2-2z^2-\frac{2}{z}) \frac{1}{(1-z)_+} + (\frac{2}{z}-1-z) \frac{1}{(1-\lambda z)_+} + \frac{1}{2} \left[\frac{1-z}{(1-\lambda z)^2} \right]_+ \right\} \quad (3.35) \end{aligned}$$

$$\begin{aligned} C_{3,q}^{(n_l)} &= \frac{4}{3} \left\{ - \left(4 + \frac{1}{2\lambda} + \frac{\pi^2}{3} + \frac{1+3\lambda}{2\lambda} K_A \right) \delta(1-z) + \right. \\ &- \frac{(1+z^2) \ln z}{1-z} + (1+z^2) \left[\frac{2 \ln(1-z) - \ln(1-\lambda z)}{1-z} \right]_+ + \\ &\left. - (1+z^2) \frac{1}{(1-z)_+} + (1-z) \frac{1}{(1-\lambda z)_+} + \frac{1}{2} \left[\frac{1-z}{(1-\lambda z)^2} \right]_+ \right\} \quad (3.36) \end{aligned}$$

where

$$\int_0^1 dz t(z) [f(z)]_+ = \int_0^1 dz (t(z) - t(1)) f(z) \quad (3.37)$$

$$\begin{aligned}
C_{1,g}^{(n_l)} &= \frac{1}{2}[z^2 + (1-z)^2] \left(\ln \frac{1-\lambda z}{(1-\lambda)z} + 2\ln(1-z) - \ln(1-\lambda z) - \ln z \right) + \\
&+ [4 - 4(1-\lambda)]z(1-z) + \frac{(1-\lambda)z}{1-\lambda z} - 1 + (1-\lambda)z \ln \frac{1-\lambda z}{(1-\lambda)z} (2 - 4\lambda z)
\end{aligned}$$

$$\begin{aligned}
C_{2,g}^{(n_l)} &= \frac{1}{2}[z^2 + (1-z)^2] \left(\ln \frac{1-\lambda z}{(1-\lambda)z} + 2\ln(1-z) - \ln(1-\lambda z) - \ln z \right) + \\
&+ [8 - 18(1-\lambda) + 12(1-\lambda)^2]z(1-z) + \frac{1-\lambda}{1-\lambda z} - 1 + \\
&+ (1-\lambda)z \ln \frac{1-\lambda z}{(1-\lambda)z} (6\lambda - 12\lambda^2 z)
\end{aligned} \tag{3.38}$$

$$\begin{aligned}
C_{3,g}^{(n_l)} &= \frac{1}{2}[z^2 + (1-z)^2] \left(\ln \frac{1-\lambda z}{(1-\lambda)z} + 2\ln(1-z) - \ln(1-\lambda z) - \ln z \right) + \\
&+ 2(1-\lambda)z(1-z) + (1-\lambda)z \ln \frac{1-\lambda z}{(1-\lambda)z} [2z(1+\lambda) - 2]
\end{aligned} \tag{3.39}$$

An example of massless limit of the x -space coefficient functions for gluon coefficient function $C_{2,g}$ reads

$$C_{2,g}^{(n_l,0)} = [z^2 + (1-z)^2] \left(\ln \frac{1-z}{z} + \frac{1}{2} \ln \frac{Q^2 + m^2}{m^2} \right) + (8z(1-z) - 1) \tag{3.40}$$

These expressions can be transformed to N -space, yielding the results given in Appendix A. Note that this is one of the key-points of the whole NNPDF treatment, at the base of the FastKernel framework and of the implementation of heavy quark effects.

Chapter 4

Benchmarks and Impact on LHC Physics

The coefficient functions presented in the previous chapter, computed in the Mellin N-space, and their massless limit have been implemented in the FastKernel framework, following the FONLL general-mass scheme. In our dataset we include all the relevant charm structure function $F_2^c(x, Q^2)$ data from the H1 and ZEUS experiments at HERA. Here below I show in Fig.4.1 the kinematical coverage, while in Table 4.2 are presented also the experimental uncertainties of these datasets.

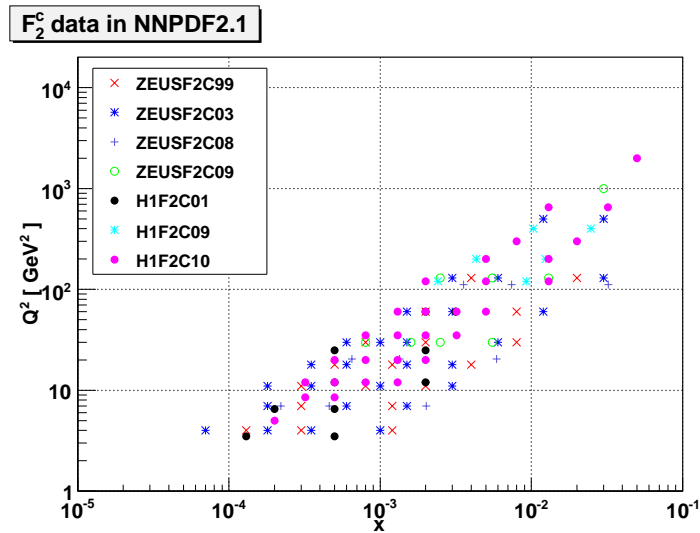


Figure 4.1: Experimental datasets for F_2^c which enter the analysis (Table 4.1). The various sets from H1 and ZEUS have been separated for illustration.

Experiment	Set	Ref.	Points	x_{\min}	x_{\max}	Q_{\min}^2	Q_{\max}^2
ZEUSF2C			69 (62)				
	ZEUSF2C99	[51]	21 (18)	$5 \cdot 10^{-5}$ ($1.3 \cdot 10^{-4}$)	0.02	1.8 (4)	130
	ZEUSF2C03	[52]	31 (27)	$3 \cdot 10^{-5}$ ($7 \cdot 10^{-5}$)	0.03	2.0 (4.0)	500
	ZEUSF2C08	[53]	9	$2.2 \cdot 10^{-4}$	0.032	7.0	112
	ZEUSF2C09	[54]	8	$8 \cdot 10^{-4}$	0.03	30	1000
H1F2C			47 (45)				
	H1F2C01	[55]	12 (10)	$5 \cdot 10^{-4}$	$3.2 \cdot 10^{-3}$	1.5	60
	H1F2C09	[56]	6	$2.4 \cdot 10^{-4}$	0.025	120	400
	H1F2C10	[57]	29	$2 \cdot 10^{-4}$	0.05	5.0	2000
Total			3554				

Table 4.1: Kinematical coverage of the various $F_2^c(x, Q^2)$ HERA data sets included in the present analysis. In parenthesis results after kinematical cuts. The last line gives the total number of datapoints included in the analysis.

Experiment	Set	Ref.	σ_{stat} (%)	σ_{sys}	σ_{norm} (%)	σ_{tot} (%)
ZEUSF2C			69 (62)			
	ZEUSF2C99	[51]	15.9	14.0	1.6	21.4
	ZEUSF2C03	[52]	11.2	14.8	2.2	19.0
	ZEUSF2C08	[53]	15.5	27.0	2.6	31.4
	ZEUSF2C09	[54]	15.9	23.3	2.6	28.5
H1F2C			47 (45)			
	H1F2C01	[55]	26.4	0.0	1.5	26.4
	H1F2C09	[56] 19.0	9.9	2.6	21.6	
	H1F2C10	[57]	8.0	12.0	2.6	14.9

Table 4.2: Average of the various experimental uncertainties of the $F_2^c(x, Q^2)$ HERA data sets included in the present analysis.

We consider for our benchmarks the F_2^c structure function. In general, a generic charm structure function can be written in the FONLL general mass scheme with damping factor (FONLL-A-Damp from now on) as

$$\begin{aligned}
F_c^{\text{FONLL}}(N, Q^2) &= F_c^{(n_i)}(N, Q^2) \\
&- \theta(Q^2 - m^2) \left(1 - \frac{m^2}{Q^2}\right)^2 \left[F_c^{(n_i, 0)}(N, Q^2) - F_h^{(n_i+1)}(N, Q^2) \right]
\end{aligned}
\tag{4.1}$$

and note that here the structure function is written in the N-space and so it contains the Mellin transforms of the massive coefficient functions given in the previous section.

It can easily be seen in Fig. 4.2 how the FONLL-A-Damp smoothly interpolates between the FFN scheme near threshold and the massless scheme at large Q^2 . This smooth matching is guaranteed by the damping factor in Eq. 4.1 which acts as a threshold suppression factor, in analogy with χ -scaling in S-ACOT- χ . Note that the PDFs and settings are identical to those of the Les Houches Heavy quark benchmark comparison.

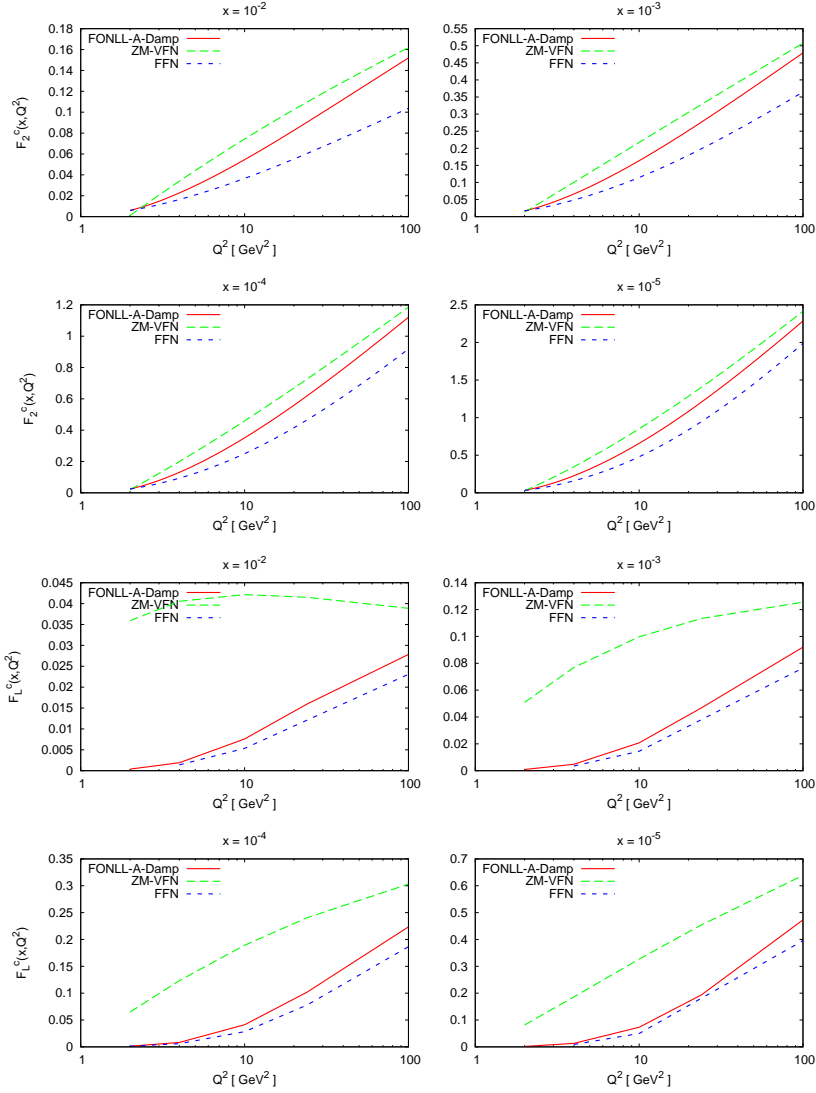


Figure 4.2: The charm structure functions $F_2^c(x, Q^2)$ and $F_L^c(x, Q^2)$ as a function of Q^2 for different values of x from $x = 10^{-5}$ to $x = 10^{-2}$ in various heavy quark schemes, computed using the FastKernel method: FONLL-A-Damp, ZM-VFN and the FFN scheme. The iterated solution of the DGLAP evolution equations has been used, and the PDFs and settings are identical to those of the Les Houches Heavy quark benchmark comparison.

The impact of heavy quark effects in DIS structure functions is further quantified in Fig. 4.3, where the relative difference between the ZM and FONLL-A-Damp schemes is computed as a function of x and Q^2 , both for the inclusive structure functions F_2^p and F_L^p and for the charm structure functions F_2^c and

x	FONLL-A-Damp			FFN		
	FONLLdis	FastKernel	Accuracy	FONLLdis	FastKernel	Accuracy
$Q^2 = 4 \text{ GeV}^2$						
10^{-5}	0.1507	0.1501	0.4%	0.1088	0.1091	0.3%
10^{-4}	0.0936	0.0931	0.5%	0.0697	0.0698	0.1%
10^{-3}	0.0506	0.0504	0.4%	0.0392	0.391	0.2%
10^{-2}	0.0174	0.0177	1.5%	0.0136	0.0137	0.7%
$Q^2 = 10 \text{ GeV}^2$						
10^{-5}	0.563	0.561	0.4%	0.3598	0.3602	0.1%
10^{-4}	0.312	0.311	0.3%	0.2007	0.2011	0.2%
10^{-3}	0.1499	0.1495	0.3%	0.0981	0.0982	0.1%
10^{-2}	0.05056	0.05052	0.1%	0.0328	0.0327	0.3%
$Q^2 = 100 \text{ GeV}^2$						
10^{-5}	2.28636	2.28577	0.02%	1.9779	1.9877	0.5%
10^{-4}	1.12186	1.12082	0.1%	0.9161	0.9184	0.3%
10^{-3}	0.48008	0.47919	0.2%	0.3644	0.3647	0.1%
10^{-2}	0.15207	0.15200	0.04%	0.1037	0.1038	0.1%

Table 4.3: Results of the benchmark comparison for the $F_{2c}(x, Q^2)$ structure function in the FONLL-A-Damp scheme for the FONLLdis code and for the FastKernel framework. Results are provided at the benchmark kinematical points in x, Q^2 . For completeness we also show the analogous results for the case of the massive scheme results.

x	FONLL-A-Damp			FFN		
	FONLLdis	FastKernel	Accuracy	FONLLdis	FastKernel	Accuracy
$Q^2 = 4 \text{ GeV}^2$						
10^{-5}	0.0130174	0.013094	0.6%	0.009077	0.009081	0.04%
10^{-4}	0.008347	0.008316	0.4%	0.005913	0.005910	0.05%
10^{-3}	0.004795	0.004778	0.3%	0.003511	0.003509	0.06%
10^{-2}	0.001910	0.001907	0.2%	0.001403	0.001406	0.2%
$Q^2 = 10 \text{ GeV}^2$						
10^{-5}	0.073235	0.073022	0.3%	0.049856	0.049982	0.3%
10^{-4}	0.041392	0.041251	0.3%	0.028402	0.028423	0.07%
10^{-3}	0.020754	0.020707	0.2%	0.014463	0.014456	0.05%
10^{-2}	0.007616	0.007595	0.3%	0.005350	0.005346	0.07%
$Q^2 = 100 \text{ GeV}^2$						
10^{-5}	0.471889	0.4729	0.2%	0.3955	0.397855	0.6%
10^{-4}	0.2236	0.2235	0.1%	0.18656	0.186914	0.2%
10^{-3}	0.0920	0.09188	0.1%	0.0765	0.076393	0.1%
10^{-2}	0.027822	0.02782	0.1%	0.023079	0.023100	0.1%

Table 4.4: Same as Table 4.3 for the $F_{Lc}(x, Q^2)$ structure function.

In Fig. 4.4 we show the comparison between important LHC observables at $\sqrt{s} = 7$ TeV: W^\pm and Z^0 production, $t\bar{t}$ production and Higgs production via gluon fusion at $m_H = 120$ GeV. We show the predictions from NNPDF2.0 and NNPDF2.1 together with those of CTEQ6.6 and MSTW08. In the latter cases we show results for at the respective default value of $\alpha_s(M_Z)$ and at the common value of $\alpha_s(M_Z) = 0.119$. All observables have been computed with MCFM, and have been tabulated as in Table 4.5. The analogous results for $\sqrt{s} = 14$ TeV are given in Fig. 4.5 and in Table 4.5.

The effect of heavy quark masses and the new F_2^c data on physical observables can be assessed by comparing the predictions for LHC observables from NNPDF2.0 and NNPDF2.1. We can see that the effect of heavy quark masses at $\sqrt{s} = 7$ TeV is always at most at the same level of the 1-sigma PDF uncertainty (for W^- and Z^0 production). A 1/2-sigma effect is observed for W^+ production, while the predictions for $t\bar{t}$ and Higgs production are essentially identical. The same can be applied for the $\sqrt{s} = 14$ TeV case, where as expected the heavy quark mass effects are slightly larger, though the predictions from NNPDF2.0 and NNPDF2.1 are always compatible at the 1-sigma level.

If we compare NNPDF2.1 with the other global PDF sets, CTEQ6.6 and MSTW08, we find that first of all in general using a common value of the strong coupling α_s leads to an improved agreement between the predictions of the various groups. At 7 TeV NNPDF2.1 is in perfect agreement with the predictions of the other sets, with only marginal agreement found though with CTEQ for Higgs production. The spread of predictions is somewhat larger at 14 TeV, where for example the agreement between NNPDF2.1 and CTEQ6.6 is only marginal for $t\bar{t}$ and W^+ production. In all cases, we are in very good agreement with the MSTW08 predictions.

	$\sigma(W^+)\text{Br}(W^+ \rightarrow l^+\nu_l)$ [nb]	$\sigma(W^+)\text{Br}(W^+ \rightarrow l^+\nu_l)$ [nb]	$\sigma(Z^0)\text{Br}(Z^0 \rightarrow l^+l^-)$ [nb]
NNPDF2.0	5.84 ± 0.14	3.97 ± 0.09	0.91 ± 0.02
NNPDF2.1	5.93 ± 0.16	4.07 ± 0.10	0.93 ± 0.02
CTEQ6.6 - $\alpha_s = 0.118$	6.05 ± 0.12	4.10 ± 0.09	0.94 ± 0.02
CTEQ6.6 - $\alpha_s = 0.119$	6.06 ± 0.12	4.11 ± 0.09	0.95 ± 0.02
MSTW08 - $\alpha_s = 0.119$	5.91 ± 0.11	4.16 ± 0.08	0.94 ± 0.02
MSTW08 - $\alpha_s = 0.120$	5.95 ± 0.11	4.19 ± 0.08	0.95 ± 0.02

	$\sigma(t\bar{t})$ [pb]	$\sigma(H, m_H = 120 \text{ GeV})$ [pb]
NNPDF2.0	168.1 ± 7.5	11.59 ± 0.22
NNPDF2.1	168 ± 7	11.56 ± 0.27
CTEQ6.6 - $\alpha_s = 0.118$	156.0 ± 6.7	10.92 ± 0.20
CTEQ6.6 - $\alpha_s = 0.119$	160.1 ± 6.7	11.07 ± 0.20
MSTW08 - $\alpha_s = 0.119$	164.4 ± 4.9	11.48 ± 0.18
MSTW08 - $\alpha_s = 0.120$	168.1 ± 4.9	11.69 ± 0.18

Table 4.5: Cross sections for W, Z, $t\bar{t}$ and Higgs production at the LHC at $\sqrt{s} = 7$ TeV and the associated PDF uncertainties. All quantities have been computed at NLO using MCFM for the NNPDF2.1, NNPDF2.0, CTEQ6.6 and MSTW08 PDF sets. All uncertainties shown are one-sigma level. See Fig. ?? for the graphical representation of the results of this table.

	$\sigma(W^+)\text{Br}(W^+ \rightarrow l^+\nu_l)$ [nb]	$\sigma(W^+)\text{Br}(W^+ \rightarrow l^+\nu_l)$ [nb]	$\sigma(Z^0)\text{Br}(Z^0 \rightarrow l^+l^-)$ [nb]
NNPDF2.0	11.59 ± 0.27	8.56 ± 0.17	1.94 ± 0.04
NNPDF2.1	11.97 ± 0.27	8.84 ± 0.19	1.99 ± 0.05
CTEQ6.6 - $\alpha_s = 0.118$	12.44 ± 0.28	9.13 ± 0.22	2.07 ± 0.05
CTEQ6.6 - $\alpha_s = 0.119$	12.49 ± 0.28	9.17 ± 0.22	2.08 ± 0.05
MSTW08 - $\alpha_s = 0.119$	11.95 ± 0.22	9.03 ± 0.17	2.01 ± 0.04
MSTW08 - $\alpha_s = 0.120$	12.06 ± 0.22	9.10 ± 0.17	2.03 ± 0.04

	$\sigma(t\bar{t})$ [pb]	$\sigma(H, m_H = 120 \text{ GeV})$ [pb]
NNPDF2.0	942 ± 21	37.3 ± 0.5
NNPDF2.1	934 ± 26	37.23 ± 0.65
CTEQ6.6 - $\alpha_s = 0.118$	871 ± 18	36.32 ± 0.82
CTEQ6.6 - $\alpha_s = 0.119$	887 ± 18	36.77 ± 0.82
MSTW08 - $\alpha_s = 0.119$	917 ± 18	37.78 ± 0.50
MSTW08 - $\alpha_s = 0.120$	934 ± 18	38.43 ± 0.50

Table 4.6: Same as Table 4.5 for the LHC at $\sqrt{s} = 14$ TeV. See Fig. ?? for the graphical representation of the results of this table.

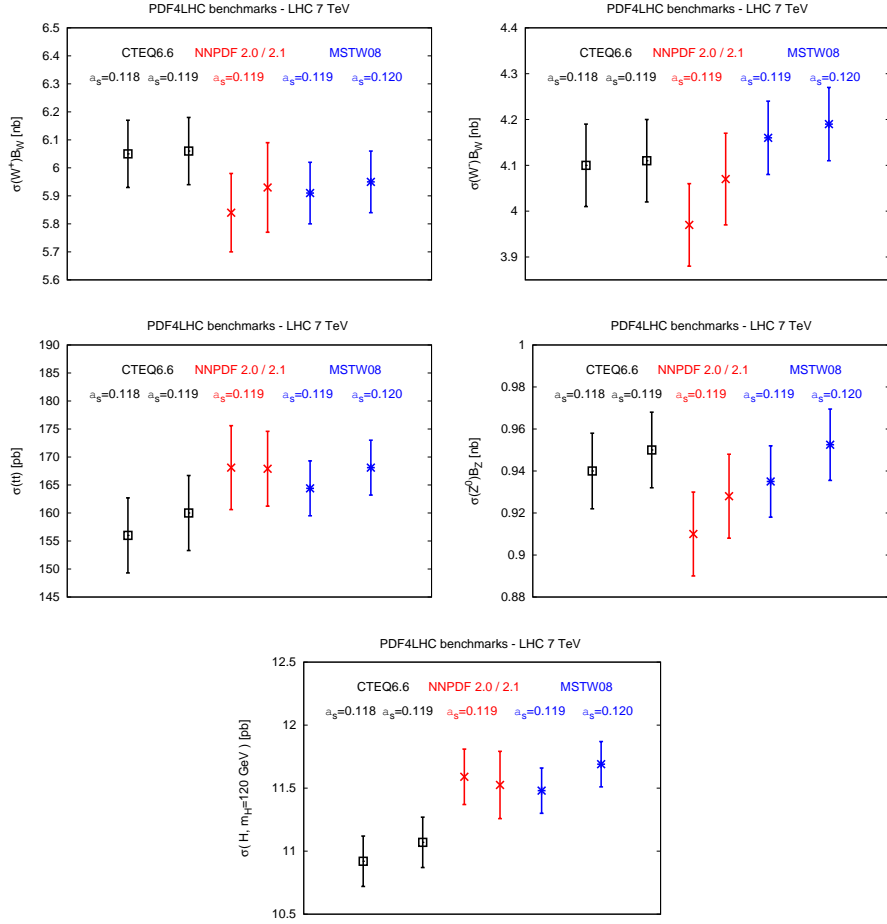


Figure 4.4: Comparison between important LHC observables at $\sqrt{s} = 7$ TeV: W^\pm and Z^0 production, $t\bar{t}$ production and Higgs production via gluon fusion at $m_H = 120$ GeV. We show the predictions from NNPDF2.0 and NNPDF2.1 together with those of CTEQ6.6 and MSTW08. In the latter cases we show results for the respective default value of $\alpha_s(M_Z)$ and at the common value of $\alpha_s(M_Z) = 0.119$. All observables have been computed with MCFM.

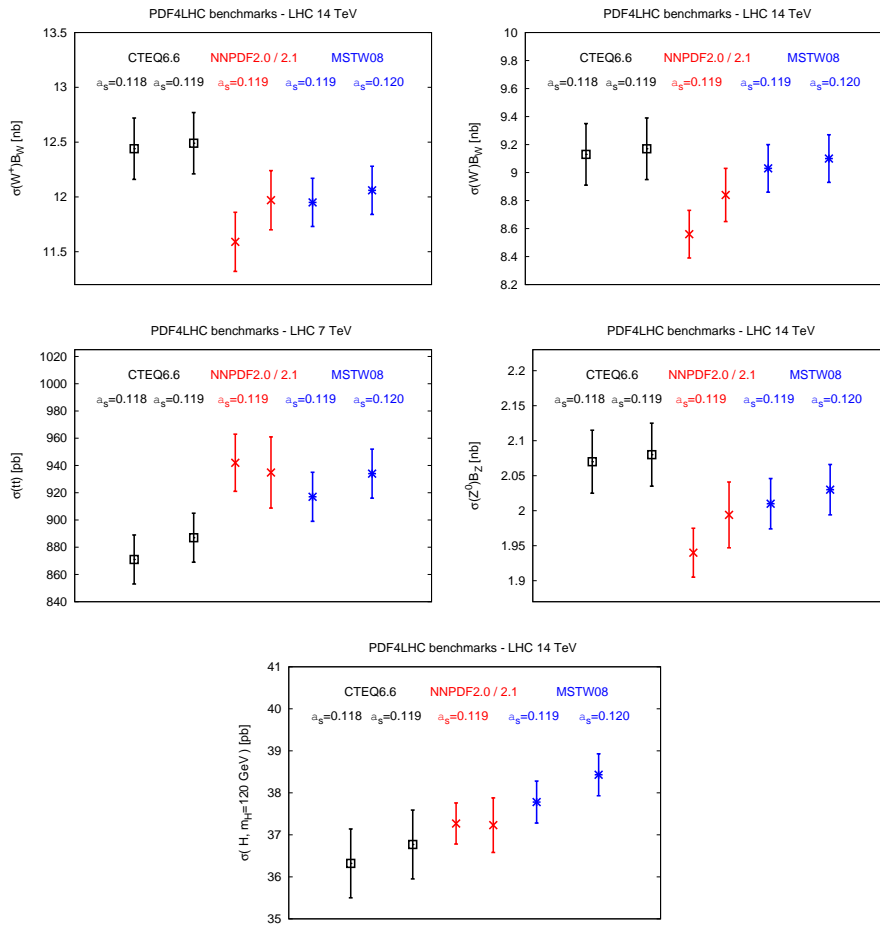


Figure 4.5: Same as Fig.4.4 at $\sqrt{s} = 14$ TeV.

Chapter 5

Conclusions and Outlook

The great improvements achieved in collider physics by the end of the last century and the great amount of data expected from the experiments displaced along the LHC ring (and that right now are collecting data) made the scientific community aware of the need for new high precision parton sets. A more rigorous and careful determination of PDFs has nowadays become compulsory.

The present work was realized using tools developed within the NNPDF Collaboration and is based on the FONLL method. A general overview of the method is given, in the innovative context of the NNPDF approach to parton sets determination. A modest effort was spent in the calculation of massive heavy quark coefficient functions in the N -space of Mellin transforms and a big effort was spent in the implementation of the method (and consequently of the coefficient function results) in the FastKernel framework, within the NNPDF Code.

Including into the analysis all the relevant charm structure function $F_2^c(x, Q^2)$ data from the H1 and ZEUS experiments at HERA, a series of checks and benchmarks are performed using the new code with heavy quark effects implemented up to $\mathcal{O}(\alpha_s)$.

Using a set of toy PDFs from the Les Houches Heavy Quark benchmark comparison, plots for the structure functions F_2^c and F_L^c are shown for different ranges of x and Q^2 (Fig. 4.2). The same functions are presented using the FastKernel method in the FONLL-A-Damp, ZMVFN and Fixed Flavour Number schemes. The behaviour of the FONLL-A-Damp scheme plot is as hoped: the curve smoothly interpolates between the FFN scheme near threshold and the massless one at high Q^2 .

In Fig. 4.3 an estimation of heavy quark effects on charm and proton structure functions is determined. In particular F_2^p can be affected up to $\sim 10\%$.

In Tabs. 4.3, 4.4 is presented a benchmark comparison for F_2^c and F_L^c where a very good accuracy is found. The results obtained with the FONLL-

A-Damp method using the FONLLdis Code (x -space) are compared with the ones obtained with the FastKernel framework (N -space): the value of accuracy stays always below the percent.

In the end also a comparison between the results from different groups for some relevant observables is shown (Tabs.4.5, 4.6 and Figs. 4.4, 4.5): the results found using the new code is in very good agreement with the ones from other groups.

In the near future the achievements of this work will be exploited to determine a first global parton fit with effects of heavy quarks included at order α_s and consequently a computation at order α_s^2 will be performed.

Appendix

Appendix A

HQ CC Massive Coefficient Functions in x and N Mellin Space

The x -space expressions given in Chapter 3 can be transformed to N -space, yielding the results presented in the following two sub-sections.

Quark Initiated Processes

$$\begin{aligned}
C_{1,q}^{(n_l)}(N, Q^2) = & \frac{4}{3} \left\{ - \left(\frac{1}{N} - \frac{1}{N+1} - 2S_1 \right) \ln \lambda - \left(4 + \frac{1}{2\lambda} + \frac{\pi^2}{3} + \frac{1+3\lambda}{2\lambda} K_A \right) + \right. \\
& \left. - 2(S_2 - \zeta_2) + \frac{1}{N^2} - \frac{1}{(N+1)^2} + \right. \\
& + \frac{1}{3} \left[\frac{3}{N^2} + \frac{3}{(N+1)^2} + 6\gamma \left(\gamma + \frac{1}{N} + \frac{1}{N+1} \right) + \pi^2 + 3 \left(S_1 - \frac{1}{N} - \gamma \right) \left(3\gamma + S_1 - \frac{1}{N} \right) + \right. \\
& \left. + 3 \left(S_1 + \frac{1}{N+1} - \gamma \right)^2 - 6 \left(\zeta_2 - S_2 + \frac{1}{N^2} \right) \right] + \\
& + J_\lambda(N) + \frac{1}{N} - \frac{1}{N+1} + 2S_1 + \\
& \left. + \frac{{}_2F_1(1, N+1, N+2; \lambda)}{N+1} - \frac{{}_2F_1(1, N+2, N+3; \lambda)}{N+2} + \frac{1}{2} \frac{{}_2F_1(2, N, N+2; \lambda)}{N(N+1)} \right\} \quad (\text{A.1})
\end{aligned}$$

where

$$J_\lambda(N) = \sum_{k=1}^{\infty} \frac{2\lambda^k}{k} \left[S_1(N+k) - S_1(k) - \frac{1}{2(N+k)(N+k+1)} \right] \quad (\text{A.2})$$

$$\begin{aligned}
C_{2,q}^{(n)}(N, Q^2) &= \frac{4}{3} \left\{ - \left(\frac{1}{N} - \frac{1}{N+1} - 2S_1 \right) \ln \lambda - \left(4 + \frac{1}{2\lambda} + \frac{\pi^2}{3} + \frac{1+3\lambda}{2\lambda} K_A \right) + \right. \\
&\quad \left. - 2(S_2 - \zeta_2) + \frac{1}{N^2} - \frac{1}{(N+1)^2} + \right. \\
&\quad \left. + \frac{1}{3} \left[\frac{3}{N^2} + \frac{3}{(N+1)^2} + 6\gamma \left(\gamma + \frac{1}{N} + \frac{1}{N+1} \right) + \pi^2 + 3 \left(S_1 - \frac{1}{N} - \gamma \right) \left(3\gamma + S_1 - \frac{1}{N} \right) + \right. \right. \\
&\quad \left. \left. + 3 \left(S_1 + \frac{1}{N+1} - \gamma \right)^2 - 6 \left(\zeta_2 - S_2 + \frac{1}{N^2} \right) \right] + \right. \\
&\quad \left. + J_\lambda(N) + K_A + 2 \frac{{}_2F_1(1, N-1, N; \lambda)}{N-1} - \frac{{}_2F_1(1, N, N+1; \lambda)}{N} - \frac{{}_2F_1(1, N+1, N+2; \lambda)}{N+1} + \right. \\
&\quad \left. + 2 \left(S_1 + \frac{1}{N+1} - \frac{1}{N-1} \right) + \frac{1}{2} \frac{{}_2F_1(2, N, N+2; \lambda)}{N(N+1)} \right\} \quad (\text{A.3})
\end{aligned}$$

$$\begin{aligned}
C_{3,q}^{(n)}(N, Q^2) &= \frac{4}{3} \left\{ - \left(\frac{1}{N} - \frac{1}{N+1} - 2S_1 \right) \ln \lambda - \left(4 + \frac{1}{2\lambda} + \frac{\pi^2}{3} + \frac{1+3\lambda}{2\lambda} K_A \right) + \right. \\
&\quad \left. - 2(S_2 - \zeta_2) + \frac{1}{N^2} - \frac{1}{(N+1)^2} + \right. \\
&\quad \left. + \frac{1}{3} \left[\frac{3}{N^2} + \frac{3}{(N+1)^2} + 6\gamma \left(\gamma + \frac{1}{N} + \frac{1}{N+1} \right) + \pi^2 + 3 \left(S_1 - \frac{1}{N} - \gamma \right) \left(3\gamma + S_1 - \frac{1}{N} \right) + \right. \right. \\
&\quad \left. \left. + 3 \left(S_1 + \frac{1}{N+1} - \gamma \right)^2 - 6 \left(\zeta_2 - S_2 + \frac{1}{N^2} \right) \right] + J_\lambda(N) + \frac{1}{N} - \frac{1}{N+1} - 2S_1 + \right. \\
&\quad \left. + \frac{{}_2F_1(1, N, N+1; \lambda)}{N} - \frac{{}_2F_1(1, N+1, N+2; \lambda)}{N+1} + \frac{1}{2} \frac{{}_2F_1(2, N, N+2; \lambda)}{N(N+1)} \right\} \quad (\text{A.4})
\end{aligned}$$

Gluon Initiated Processes

$$\begin{aligned}
C_{1,g}^{(n_l)}(N, Q^2) &= -\frac{1}{2} \left[\frac{2}{N+2} - \frac{2}{N+1} + \frac{1}{N} \right] \ln \lambda + \\
&+ \frac{4 - 2(N-3)N - N(N^2 + N + 2)(\ln(1-\lambda) + 2S_1)}{2N^2(N+1)(N+2)} + \\
&+ \left(4 - 4(1-\lambda) \right) \left(\frac{1}{N+1} - \frac{1}{N+2} \right) + (1-\lambda) \frac{{}_2F_1(1, N+1, N+2, \lambda)}{N+1} - \frac{1}{N} + \\
&+ 2(1-\lambda) \frac{{}_2F_1(1, N+1, N+2, \lambda)}{(N+1)^2} - 4(1-\lambda) \frac{{}_2F_1(1, N+1, N+2, \lambda) - 1}{(N^2 + 3N + 2)}
\end{aligned} \tag{A.5}$$

$$\begin{aligned}
C_{2,g}^{(n_l)}(N, Q^2) &= -\frac{1}{2} \left[\frac{2}{N+2} - \frac{2}{N+1} + \frac{1}{N} \right] \ln \lambda + \\
&+ \frac{4 - 2(N-3)N - N(N^2 + N + 2)(\ln(1-\lambda) + 2S_1)}{2N^2(N+1)(N+2)} + \\
&+ \frac{2 - 6\lambda + 12\lambda^2}{N^2 + 3N + 2} + (1-\lambda) \frac{{}_2F_1(1, N, N+1, \lambda)}{N} - \frac{1}{N} + \\
&+ 6\lambda(1-\lambda) \frac{{}_2F_1(1, N+1, N+2, \lambda)}{(N+1)^2} - 12\lambda(1-\lambda) \frac{{}_2F_1(1, N+1, N+2, \lambda) - 1}{(N^2 + 3N + 2)}
\end{aligned} \tag{A.6}$$

$$\begin{aligned}
C_{3,g}^{(n_l)}(N, Q^2) &= -\frac{1}{2} \left[\frac{2}{N+2} - \frac{2}{N+1} + \frac{1}{N} \right] \ln \lambda + \\
&+ \frac{4 - 2(N-3)N - N(N^2 + N + 2)(\ln(1-\lambda) + 2S_1)}{2N^2(N+1)(N+2)} + 2(1-\lambda) \left(\frac{1}{N+1} - \frac{1}{N+2} \right) + \\
&- 2(1-\lambda) \frac{{}_2F_1(1, N+1, N+2, \lambda)}{(N+1)^2} + 2(1-\lambda^2) \frac{{}_2F_1(1, N+1, N+2, \lambda) - 1}{\lambda(N^2 + 3N + 2)}
\end{aligned} \tag{A.7}$$

Bibliography

- [1] D.J. Gross, F. Wilczek, Phys. Rev. D8, (1974) 3633.
- [2] D.J. Gross, F. Wilczek, Phys. Rev. D9, (1974) 980.
- [3] R. McElhaney and S.F. Tuan, Phys. Rev. D8 (1973) 2267.
- [4] T. Kawaguchi and H. Nakkagawa, Analysis of Scaling Violation in Terms of Theories with Anomalous Dimensions, KUNS 380, 1976.
- [5] A. De Rujula, H. Georgi and H.D. Politzer, Ann. Phys. 103 (1977) 315.
- [6] P.W. Johnson and W.k. Tung, Nucl. Phys. B121 (1977) 270.
- [7] M. Gluck and E. Reya, Phys. Rev. D14 (1976) 3034.
- [8] I. Hinchliffe and C.H. Llewellyn Smith, Nucl. Phys. B128 (1977) 93.
- [9] M. Gluck, E. Hoffmann and E. Reya, Zeit. Phys. C13 (1982) 119.
- [10] D.W. Duke and J.F. Owens, Phys. Rev. D30 (1984) 49.
- [11] E. Eichten et al., Rev. Mod. Phys. 56 (1984) 579.
- [12] A. Devoto et al., Phys. Rev. D27 (1983) 508.
- [13] R.G. Roberts, The Structure of the proton: Deep inelastic scattering (Cambridge University Press, 1990).
- [14] R.K. Ellis, W.J. Stirling and B.R. Webber, QCD and collider physics (Cambridge University Press, 1996).
- [15] HEPDATA, <http://durpdg.dur.ac.uk/hepdata/pdf.html>.
- [16] M. Klein and R. Yoshida, (2008), 0805.3334.
- [17] CTEQ, H.L. Lai et al., Eur. Phys. J. C12 (2000) 375, hep-ph/9903282.
- [18] A.D. Martin et al., Eur. Phys. J. C23 (2002) 73, hep-ph/0110215.
- [19] M. Diemoz et al., Z. Phys. C39 (1988) 21.

- [20] W.T. Giele and S. Keller, Phys. Rev. D58 (1998) 094023, hep-ph/9803393.
- [21] W.T. Giele, S.A. Keller and D.A. Kosower, (2001), hep-ph/0104052.
- [22] NNPDF, R.D. Ball et al., Nucl. Phys. B809 (2009) 1, 0808.1231.
- [23] LHAPDF, <http://projects.hepforge.org/lhapdf/>.
- [24] D. Bourilkov, R.C. Group and M.R. Whalley, (2006), hep-ph/0605240.
- [25] NNPDF, J. Rojo et al., (2008), 0811.2288.
- [26] The NNPDF, R.D. Ball et al., Nucl. Phys. B823 (2009) 195, 0906.1958.
- [27] M. Dittmar et al., (2009), 0901.2504.
- [28] S. Forte et al., JHEP 05 (2002) 062, hep-ph/0204232.
- [29] NNPDF, L. Del Debbio et al., JHEP 03 (2005) 080, hep-ph/0501067.
- [30] NNPDF, L. Del Debbio et al., JHEP 03 (2007) 039, hep-ph/0701127.
- [31] P.M. Nadolsky et al., Phys. Rev. D78 (2008) 013004, 0802.0007.
- [32] A.D. Martin et al., Eur. Phys. J. C63 (2009) 189, 0901.0002.
- [33] H1 and ZEUS, A.F. D et al., (2009), 0911.0884.
- [34] S. Alekhin, K. Melnikov and F. Petriello, Phys. Rev. D74 (2006) 054033, hep-ph/0606237.
- [35] T. Carli, G.P. Salam and F. Siegert, (2005), hep-ph/0510324.
- [36] M. Dittmar et al., (2005), hep-ph/0511119.
- [37] W. K. Tung, H. L. Lai, A. Belyaev, J. Pumplin, D. Stump and C. P. Yuan, JHEP 0702 (2007) 053, hep-ph/0611254.
- [38] J. C. Collins, Phys. Rev. D 58 (1998) 094002, hep-ph/9806259.
- [39] M. A. G. Aivazis, J. C. Collins, F. I. Olness and W. K. Tung, Phys. Rev. D 50 (1994) 3102, hep-ph/9312319.
- [40] J. C. Collins, F. Wilczek and A. Zee, Phys. Rev. D 18 (1978) 242.
- [41] M. 1. Kramer, F. I. Olness and D. E. Soper, Phys. Rev. D 62 (2000) 096007, hep-ph/0003035.
- [42] W. K. Tung, S. Kretzer and C. Schmidt, J. Phys. G 28 (2002) 983, hep-ph/0110247.
- [43] S. Kretzer, H. L. Lai, F. I. Olness and W. K. Tung, Phys. Rev. D 69 (2004) 114005, hep-ph/0307022.

- [44] R. S. Thorne and R. G. Roberts, Phys. Rev. D 57 (1998) 6871, hep-ph/9709442.
- [45] R. S. Thorne and R. G. Roberts, Phys. Lett. B 421 (1998) 303, hep-ph/9711223.
- [46] R. S. Thorne, Phys. Rev. D 73 (2006) 054019, hep-ph/0601245.
- [47] R. S. Thorne and W. K. Tung, hep-ph/0809.0714.
- [48] F. Olness and I. Schienbein, Nucl. Phys. Proc. Suppl. 191 (2009) 44, hep-ph/0812.3371.
- [49] M. Cacciari, M. Greco and P. Nason, JHEP 9805 (1998) 007, hep-ph/9803400.
- [50] M. Glck, S. Kretzer and E. Reya, Phys.Lett. B 380 (1996) 171-176, hep-ph/9603304.
- [51] ZEUS, J. Breitweg et al., Eur. Phys. J. C12 (2000) 35, hep-ex/9908012.
- [52] ZEUS, S. Chekanov et al., Phys. Rev. D69 (2004) 012004, hep-ex/0308068.
- [53] ZEUS, S. Chekanov et al., Eur. Phys. J. C63 (2009) 171, 0812.3775.
- [54] ZEUS, S. Chekanov et al., Eur. Phys. J. C65 (2010) 65, 0904.3487.
- [55] H1, C. Adloff et al., Phys. Lett. B528 (2002) 199, hep-ex/0108039.
- [56] H1, F.D. Aaron et al., Phys. Lett. B686 (2010) 91, 0911.3989.
- [57] H1, F.D. Aaron et al., Eur. Phys. J. C65 (2010) 89, 0907.2643.